A MONTE CARLO-BASED SIMULATION STUDY FOR ASSESSING RADIATION-INDUCED DNA DAMAGE AND CELL SURVIVAL

A Dissertation
Presented to
The Academic Faculty

By

Brian H. Lee

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Nuclear and Radiological Engineering

Georgia Institute of Technology

August 2017

A MONTE CARLO-BASED SIMULATION STUDY FOR ASSESSING RADIATION-INDUCED DNA DAMAGE AND CELL SURVIVAL

Approved by:

Dr. Chris Wang, Advisor
Department of Nuclear and Radiological Engineering
Georgia Institute of Technology

Dr. Eric Elder
Department of Nuclear and Radiological Engineering
Georgia Institute of Technology

Dr. Nolan Hertel
Department of Nuclear and Radiological Engineering
Georgia Institute of Technology

Dr. Yuhong Fan
School of Biological Sciences
Georgia Institute of Technology

Dr. William Dynan
Departments of Radiation Oncology and Biochemistry
Emory University

Date Approved: July 20, 2017

To Gloria and Esther

ACKNOWLEDGEMENTS

It is my greatest honor to thank and acknowledge my advisor, Dr. Chris Wang, for his continued assistance and guidance throughout my time at Georgia Tech. His insight in radiobiology and mentorship have been instrumental in my development and confidence as a researcher, and his display of character has helped in my own personal maturation.

I am tremendously thankful to Dr. Eric Elder for the many times he has opened doors for me professionally in my clinical physics educational, as well as in helping springboard my efforts in this research work. I would like to express my gratitude to Dr. Nolan Hertel for his constant encouragement, helpful feedback, and for bringing out the best of my engineering roots.

To Dr. William Dynan I am grateful for expanding my aspirations in research, in challenging me to seek answers for scientific progress and in expanding my network of radiobiologists. To Dr. Yuhong Fan I am grateful for expanding my awareness of how remarkable and nuanced the biological sciences are, and for elevating my amazement of the intricacies of DNA.

I am grateful to Gocha, Mahesh, Jane, Dan, Hualin, Seelan, Plato and Zach at Northwestern Memorial Hospital for their helping me understand the gravity of radiation oncology throughout my residency training.

To my parents, Kyungsub and Bongsoo Lee, I could never fully express my gratitude for all of their love and sacrifice. My sisters, Eunice and Jane, are my inspiration for always putting forth my best effort. My daughter, Esther Dahyoung Lee, has been my strongest motivation to complete this task. To Gloria Eunyoung Lee, the love of my life and my source of financial stability, thanks for everything.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

This dissertation describes a Monte Carlo-based simulation study that integrates charged particle track structures and cell nucleus DNA organizations to quantify DNA and chromatin damage as well as the cell survival rate for various radiation types. Geant4-DNA, a detailed Monte Carlo code for particle track simulation at the nanometer scale, was employed for the production of charged particle tracks. The cell nucleus DNA organizations modeled in the study include chromatin domains, chromatin fibers, and chromosome territories. The positioning and orientation of these organizations in a cell nucleus are based on the Monte Carlo method. This study also includes a stochastic method for simulating the production of DNA double strand breaks (DSBs) and DSB misrejoining events, which can be used to generate chromosome aberrations and cell survival curves.

In the presented work we are able to characterize differences in the spatial distribution pattern of DSBs produced by low-LET electrons, ultrasoft X-rays, protons, helium ions, and carbon ions. A core element of this Monte Carlo study is that subtle nuances of charged particle interactions and DNA damage are retained. The results include the unique spatial distributions of nanometer scale clusters of energy deposition events as well as the spatial distribution of DSBs for the different radiation types. The spatial distribution of DSBs, in turn, allows the estimate of the number of potential DSB misrejoining events, Chromosome aberrations, and cell survival probability. The stochastic nature of the simulation method allows the cell survival fraction to be estimated on the cell-by-cell basis, reflecting the true nature of radiation-induced cell killing effect.

In the presented work we also show that the new radiobiological model may find applications in radiotherapy and radiation protection. In radiotherapy, it can be used to estimate the RBE values for radiotherapy that employs radiation types other than the conventional X-rays (e.g. protons, neutrons, and carbon ions). In radiation protection, it can be used to estimate the radiation weighting factors for the various radiation types.

CHAPTER 1
INTRODUCTION

The present status of radiobiology as a field of study has developed greatly since the initial discoveries of ionizing radiation in the late 1800's[1]. Owing to its dependence on advanced principles of biological systems and subcellular response [2] as well as physical behaviors of charged particle interactions[3, 4, 5], radiobiology is by nature a collaborative research exploration involving the input of specialized expertise from multiple disciplines. While the current state of the art has consistently advanced in each of the respective disciplines of clinical physics practice and experimental radiobiology[6, 7, 8, 9], at present the collaboration between the two disciplines has not fully caught up to the most recent tools and discoveries that both have to offer.

With respect to clinical physics practice, the primary responsibility is in the safe and effective delivery of therapeutic doses of radiation to patients under treatment for various malignancies and functional impairments [10]. Much of the emphasis in clinical physics practice typically, and understandably, focuses on the most direct benefit to patients given the practical necessities of consistent and ethical healthcare delivery. While the field has developed over time, more of the focus of physics research has been on image guided therapy for target localization [11, 12], dose calculation engines [13] and machine characterization for accurate simulation of radiation delivery [14, 15], and standardization of practice in the form of IAEA and AAPM Task Group reports [16]. Owing to the typical workload of an academic physicist, many opt to specialize in research with direct clinical implementations. However, in the past decade several tools have been developed and made available that are very useful for application in more academic and theoretical simulations for quantifying the effect of radiation damage at the DNA scale [17, 6, 5, 7, 18].

Regarding advances in biology at this point in history, it would be no exaggeration

to say that headline developments are reported in front page news articles on at least a bi-monthly basis. It has been around two decades since the human genome project was completed, and the impact on cancer research has been substantial, with genetic testing being a standard of care for determining optimal treatment [19, 20]. Within the field of radiobiology itself, the cataloguing of response pathways to ionizing radiation is regularly published upon and widely reported. Even the very organization of DNA itself, not only in how it is expressed but also in how it is organized, is a specialized field that itself has developed and expanded over the past decade[21, 22, 23, 24].

The aim of the presented work has been to narrow the gaps between clinical physics and experimental biology for the advancement of radiobiology[18]. For this goal we present a Monte Carlo-based biophysical model that is up-to-date with both the structural hierarchy of DNA organization in a cell nucleus as well as in the available charged particle track simulations [5, 22, 21]. Chapter 2 provides the background knowledge needed for the Monte Carlo simulation study. Chapter 3 describes the overall structure of the Monte Carlo-based study. The remaining chapters respectively describe the details of the four parts of the simulation study: development of charged particle track structure data library (Chapter 4), simulation of the cell nucleus (Chapter 5), production of DNA double strand breaks (Chapter 6), and evaluation of cell survival (Chapter 7). The results and discussion on radiation-induced DNA double strand breaks and cell death are included in Chapters 6 and 7. Chapter 8 gives the conclusions.

CHAPTER 2
BACKGROUND

The purpose of the presented work is to develop a self-consistent tool that can quantify the differences in expected DNA damage to a comprehensive variety of exposures to ionizing radiation, irrespective of radiation type, energy, dose quantity, dose rate, or any combination of multiple forms therein. It is a daunting task, given that clinical situations typically involve variations of multiple parameters, and adjustment of multiple variables can lead to results that are unexpected [25, 26, 27, 28]. One of the novelties to the presented work is that it is based on both the Monte Carlo simulation of charged particle tracks and the biological simulation of subnuclear DNA structures.

2.1 Basics of Radiobiology

Prior to delving into the detailed background underlying the model, the overarching principles behind the model require some explaining to more simply navigate the presented work. The first principle is that the particle tracks have been compiled using Monte Carlo simulations, widely understood to be the gold standard for both general radiation transport as used in the nuclear sciences and for clinical treatment planning [29, 30, 31, 3, 4]. Much of the discussion into the details of charged particle tracks and their influence on DNA damage is derived from observations taken from the Monte Carlo codes, which themselves have been verified through experimental results over a long history[32, 33, 34, 35, 36, 37, 38, 8]. The second principle is that specific structural features of DNA organization in the cell nucleus is the modeled target for radiation damage in our biophysical model[21, 22, 39, 40]. A third principle tied to the second principle is that the assessment of cellular damage, misrejoining of DNA double strand breaks, has a dependence on the likelihood of interaction of DNA double strand breaks based on their structural origin[41, 42, 43]. Almost

all of the description within this body of work is intended to explain how the components work together, justify the parameters selected for the simulation, and present some of the findings of the work itself.

2.1.1 Cell Survival and the Linear Quadratic Equation

Cell survival curves are common practice for reporting the efficacy of radiation to induce cell death for controlled variables[1, 10, 9]. The relationship between dose and the surviving fraction of cells exposed to radiation in vitro are necessary for the testing of a multitude of conditions, including dose rate, cell lines, radiation types, and radiation energy[9, 44, 45]. Furthermore, unexpected differences in response are observed through such tests, and failure of a specific simulation or radiobiological model to match experimental data indicate specific shortcomings in the model to account for the actual behavior of cells in response to radiation. While cellular pathways are complex and not fully understood, the surviving fraction of cells exposed to radiation is a straightforward method for quantifying damage even if the underlying reasons may still be under investigation.

The linear quadratic (LQ) equation is a deterministic approach used in radiobiology to plot cell survival versus dose delivered and is shown in Fig. 2.1. The cell survival S is:

$$S = e^{-(\alpha D + \beta D^2)} \tag{2.1}$$

where D is the dose, α is the linear portion of the curve corresponding, while β is the quadratic portion of the curve related to D^2 corresponding to the interactions between tracks [1]. α is noted to be the initial slope of the curve and is relevant to the single track effect which will be discussed in more detail, while β is more involved in interactions between different ionizing radiation tracks in a cell nucleus.

The Linear Quadratic formalism is a straightforward observation of cell survival curves, in that, within a reasonable dose range, the semi-log plot behavior of the surviving fraction starts at an initial linear slope but progresses to a quadratic curvature as the dose in-

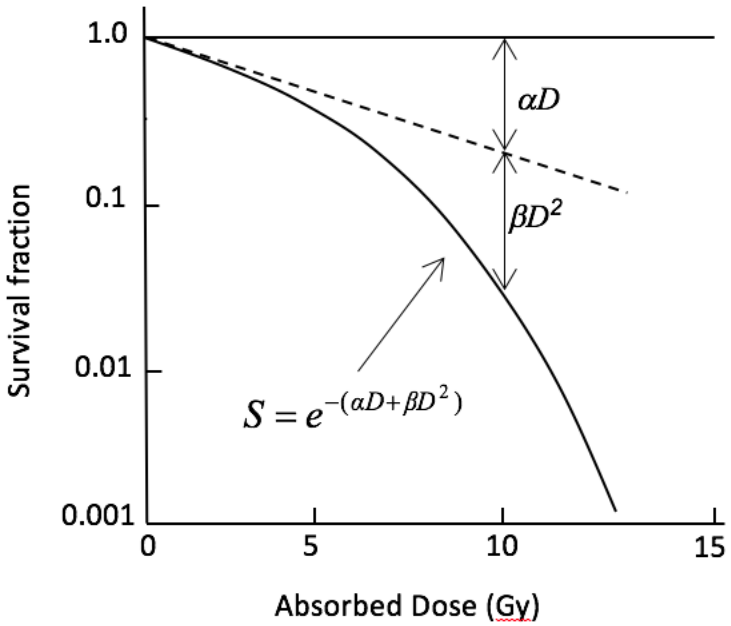


Figure 2.1: The cell survival curve described by the linear quadratic equation

creases[45, 46]. It is quick and easy to understand the LQ model, but observations of actual cell survival as the dose fraction increases indicates a more complex behavior of cellular response, as the slope typically flattens out into a linear behavior again as dose exceeds a certain threshold. Several theories exist regarding the underlying principles behind the non-adherence to pure linear quadratic behavior of cell survival curves [47], but it can be noted that to prove them out through a biophysical model allows for a more specific understanding of how damage influences the cellular response.

2.1.2 Biologically Equivalent Dose and Relative Biological Effectiveness

Charged particles are, in the simplest terms, understood to cause direct and indirect damage to DNA [1, 10, 9]. For most clinical circumstances a fairly simple application of radiobiological modeling is required, but more of the advanced current explorations across the radiological sciences require a more complete understanding of why differences in radiation type cause different cellular responses. From a nanodosimetric perspective, the most simple answer is that the charged particle track structure varies greatly between different ra-

diation types, and without a better understanding of the intricacies of the potential damage, there will be no satisfying answer as to the observed damage.

$$RBE = \frac{D_{X-ray}}{D_{test}} \tag{2.2}$$

where D_{test} is the dose of the test radiation in question required to elicit a specific biological response and $D_{X\text{-}ray}$ is the dose of X-ray radiation required to elicit the same biological response as the test radiation [48]. The assumption in the RBE equation is that the X-ray dose is in the numerator because it is less effective to induce a cellular response than the test radiation type, and in many cases it is given as the amount of radiation necessary to yield a 10% cell survival fraction.

One of the contested topics in particle therapy is in the concept of relative biological effectiveness (RBE) and biologically equivalent dose (BED) [49, 50, 51, 46]. With respect to RBE, there are several options in terms of the types of particles that can be used, but patient treatment requires some standardization of clinical prescription. The extra cost of heavy ion therapy along with limitations in dose delivery makes assumptions for such treatments challenging to work with. In addition, it is also understood that the RBE of particles is not a static value based on particle type, but it varies depending on the energy of the particles, and is not a fixed value even among cell lines [52, 53, 54]. Further phenomena that make RBE challenging to work with is that varying the radiation type and parameters will inherently change the damage pattern, which itself influences on other cellular conditions such as oxygen enhancement and hypoxic conditions. Consequences of uncertainties with varying irradiation schemes have led to unintended secondary effects[55]. This variation in RBE is one of the most pressing issues that necessitates more accurate biophysical models.

Radiobiological models become critical in their implementation particularly when crossing data across radiation types. For photon and electron radiation, this effect is not as pronounced as the RBE is consistent in the clinical treatment range. However, photons and electrons are important for tuning a model to observed cell survival. An important criteria

for a comprehensive biophysical model is that it should be accurate even for a large breadth of experimental conditions without too many fudge factors.

2.2 Various types of ionizing radiation

The RBE is known to be closely related to the linear energy transfer (LET) of the radiation. The LET is also known as the collision stopping power, which is defined as the average rate of energy loss by a charged particle (in $\mathrm{keV}/\mu\mathrm{m}$) in a medium via both the inelastic collisions (i.e. ionizations and excitations) and the elastic collisions with the atoms. The RBE, in general, is high for high-LET particles and low (i.e. approaches 1.0) for low-LET particles. For example, the RBE is about 5 for 1-MeV protons of which the LET is approximately $30\,\mathrm{keV}/\mu\mathrm{m}$ and is about 1.1 for 100-MeV protons of which the LET is approximately $0.7\,\mathrm{keV}/\mu\mathrm{m}$. The RBE-vs-LET relationship, however, is not a one-to-one relationship. It also depends on the charged particle type. For example, as shown in Figure 2.2, for the same LET of $35\,\mathrm{keV}/\mu\mathrm{m}$ the RBE of protons is roughly twice that of helium-4 ions. To be able to predict the RBE for various particle types of various LETs, one must explore the details of particle track structures. Figure 2.3 gives a visual comparison of two particle tracks having the same LET of $35\,\mathrm{keV}/\mu\mathrm{m}$, i.e. a 0.7-MeV proton and a 20-MeV helium-4 ion. They are produced by the Monte Carlo code Geant4-DNA [6, 30]. Each dot along the particle track represents an ionization or excitation event. As shown, the ionization and excitation events of the 0.7-MeV proton track are more closely spaced than that of the 20-MeV helium-4 ion track. The more closely spaced ionization and excitation events, in turn, may produce more closely spaced DNA double strand breaks (DSBs), and thus, lead to a higher chance of chromosome aberrations and cell death.

There are several conditions that can shed some light into how to better understand the challenges track structure. The first consideration to note is that track structure and dose should be considered together. While there are variations in the density of charged particle tracks, results reported with respect to RBE are in consideration with dose[1]. Statistical

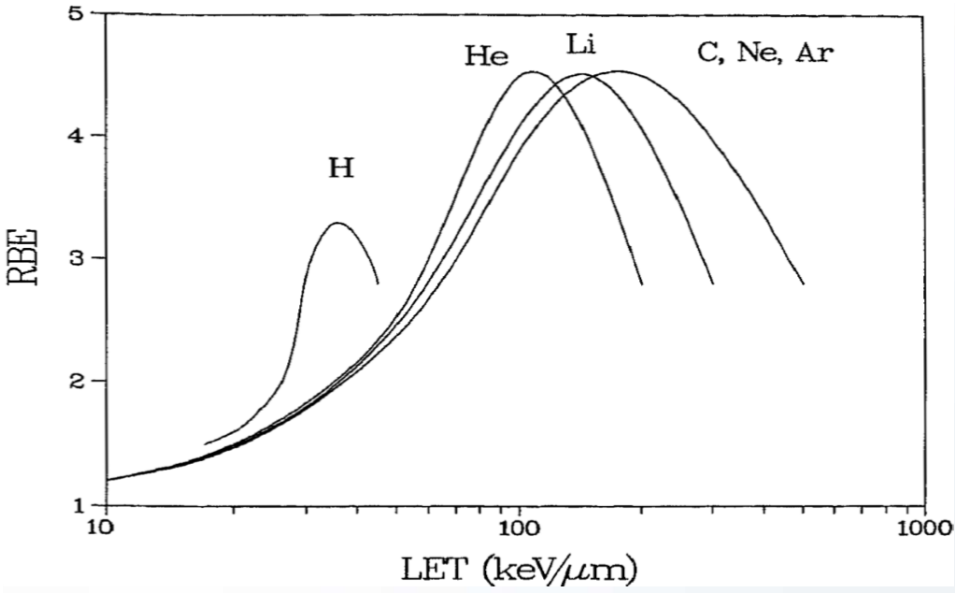
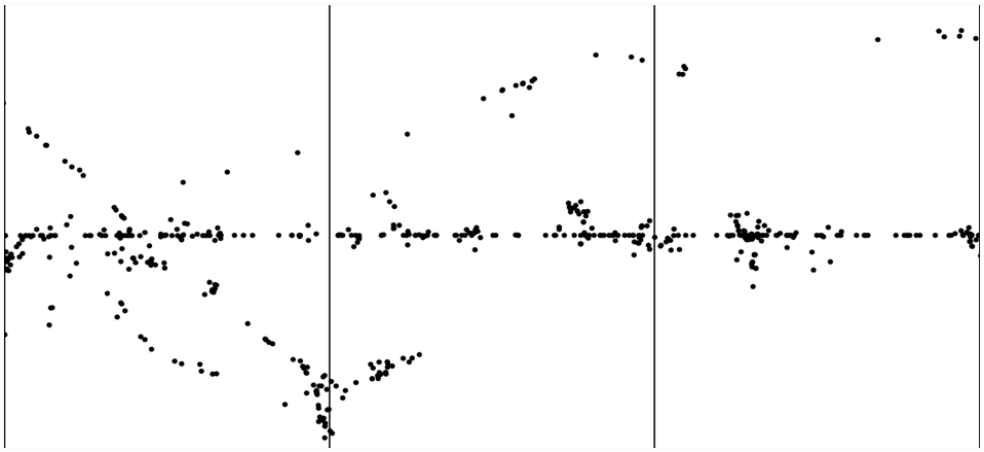


Figure 2.2: LET vs RBE of various particles based on approximations[56]

(a) 0.7 MeV proton track, 35 keV/um LET



(b) 20 MeV helium ion track, 35 keV/um LET

Figure 2.3: Geant4-DNA particle track visualization of proton and helium ion tracks of similar LET, 100nm grid

variations of dose deposition to cells increase as the LET of a particle increases, owing to a lower number of particles required for delivering the same dose[57]. With increasing LET, as well, there is typically a shorter total path length of charged particles in the cell to deposit the same dose, which when applying a DNA target model, means a smaller quantity of DNA targets will be hit for the same dose of radiation by higher LET particles.

In addition, same LET and dose does not mean the same damage when different particle types are used[56]. Due to the differences in the interaction cross sections of a medium such as water or tissue for accelerated particles, there really is no direct apples-to-apples comparison when considering different types of charged particles[52]. This in effect leads to substantially different distributions, quantities, and severities of damage dealt to the DNA structures by charged particles of seemingly similar properties. This topic is heavily contested, and understandably so, as much effort goes into developing simplified explanations of complex phenomena. While the radiological sciences would benefit from more creative biophysical modeling, experiments on various topics still need to be conducted. Furthermore, major questions can still be answered even without a comprehensive simulation in place, although a better model would greatly help investigators.

2.2.1 Electrons and photons, sparsely ionizing radiation types

The most common source of clinical and experimental data is from electron and photon irradiation at energies high enough to adequately dose a target region. These are important in radiobiological modeling as higher energy photon irradiation are considered to be the reference radiation type: typical RBE considerations compare alternative charged particles to either Cs-137 or Co-60, which feature a monoenergetic photon of 662 keV and an average photon energy of 1.25 MeV, respectively, and are assigned an RBE of 1 [57, 58]. The response of cells to such ionizing radiation types is very well documented, and covers a breadth of conditions across multiple cell lines, different dose rates, dose fractionation, oxygenation, and other considerations [59, 1].

For clinical and radiation protection purposes, all such radiation types emanating from a linear particle accelerator, or linac, are considered identically as having an RBE of 1 [60, 61, 1]. The photons produced from a linac, or even from a CT scanner or other diagnostic X-ray generators, are polyenergetic X-ray sources with peak energies typically on the order of 80 kVp to 18 MV, although a larger range is possible [62, 63]. Electron beams produced in a linac is typically defined as the patient entrance energy of the electron beam and is more typically on the order of 4 MeV to 22 MeV, although other energies are attainable [62]. In all cases, though, radiation oncologists and clinical physics personnel will treat all such therapeutic radiation beams as having an RBE of 1 [1]. In some cases, observed biological effects, such as dose rate in total body irradiation (TBI) treatments, are considered, but only insofar as an upper boundary of dose rate is maintained so as to prevent specific biological effects[64]. In a typical clinic where photons or electrons are the only sources of ionizing radiation, which is probably on the order of 99 of every 100 radiation oncology centers in the United States and internationally, the RBE is assumed to be 1.

Where clinical photon and electron beams can benefit from a more advanced radiobiological model is in the more common use of hypofractionation in clinical situations. Stereotactic radiosurgery (SRS), stereotactic radiotherapy (SRT), stereotactic body radiotherapy (SBRT), and high-dose-rate (HDR) brachytherapy are being more rapidly and widely deployed across clinics in the United States [65, 66, 67, 68]. Several published guidelines are in place for dose restrictions to critical organs that have been in widespread use to guide prescriptions, and are part of the standard curriculum of clinical physics training for some time now [65]. In addition, clinical studies on the efficacy of hypofractionation have highlighted the benefits of use [69, 70]. However, at such high doses, some of the simple assumptions about radiobiology and biologically equivalent dose (BED) break down[46]. While this may not necessitate biophysical models to prove out the efficacy of treatments currently in use, in silico simulations do have advantages of providing insight into observed phenomena and can help guide treatment decisions when linked to clinical data. Moreover,

the dose limitations are understood to be conservative, and in special cases, treatment volumes often encroach on critical organs: it would be helpful to know if there is an actual upper limit on dose delivery as patients benefiting from hypofractionation often are in poor performance status and can be more effectively treated with higher doses [70].

The main imperative of clinical physics practice in radiobiology is in accurate dose delivery such that a desired prescription matches the executed treatment [71]. Rather than worrying about how the radiation interacts at the nanoscale level of track structure, there are a broad spectrum of considerations involved in patient treatment that leads to serious uncertainties about the dose delivery to the patient's clinical treatment volume and their organs at risk [15, 16].

It is important to give credit to therapy physicists for the breadth of responsibilities they undertake; they are largely responsible for each step of treatment execution to the patient, which includes proper function of the equipment, the effective education and training of all personnel, accuracy of the beam model for each component of every radiation source in their clinic, and research and teaching considerations for the academic credibility of their institution. At the same time, the behavior of radiation, even at the nanoscale level of DNA, is most likely to be within the interest and expertise of the physicist in an academic radiation oncology team. Without the extra effort of the clinical physicist, there are few others who are as professionally and emotionally invested in the further development of this small specialty.

2.2.2 Protons

Proton therapy is an area of increased popularity and controversy in healthcare in the United States. There are good reasons for both optimism and skepticism for proton therapy: it offers the ability to deliver a therapeutic dose to a tumor with substantially lower dose delivery to normal tissues [49], but it is considerably more expensive than standard photon and electron therapies and it has its own radiobiological uncertainties that have produced

undesirable acute effects in several studies. In practical terms, it cannot be overstated on how challenging research into this area must be. The very delivery of therapeutic proton beams is a specialization that many excellent physicists have difficulty in comprehending. Most quick treatment planning introductions involve the stacking of Bragg peaks together to produce a spread out Bragg peak (SOBP) for clinical treatment. Precise tissue compensators are milled using computer numeric control (CNC) machines that undergo their own QA process.

The clinical RBE of protons is clinically accepted to be 1.1, but this is hardly a consensus number [49, 53]. Experimental cell survival curves have demonstrated highly variable RBE based on the energy of the proton beam. In addition, very little data exists for the RBE of a proton beam at all of the beam energies within the patient, from the entrance region to the range straggling portion of the SOBP. Published cell survival typically indicates the RBE of proton beams at entrance regions to match standard photon radiation with an RBE of 1. Some uncertainty exists as to where the RBE elevates, but most assume this to be a factor of the increasing dose deposition at the falling end of the Bragg peak. The RBE of 1.1 is more of an observed cell survival that involves the resulting combination of varying RBE effects of the individual protons and secondary electrons contributing to the aggregate dose in the treatment volume.

More recent publications indicate a more substantial uncertainty, especially at the distal end of the SOBP, where range straggling effects are observed. As much of the promise for proton therapy has been in pediatrics, acute effects at the distal end of a SOBP, which is commonly in healthy tissue and sometimes abutting an organ-at-risk, can be devastating. While each treatment modality will have some drawbacks, understanding the source of uncertainties can help mitigate the observed effect. In standard photon therapy, for example, the electron density of patients allows for more accurate dose calculation engines to be used to compensate for tissue heterogeneities, but significant CT artifacts introduced by metal implants and prostheses for some patients can cause such corrections to fail. Reverting

to older methods of homogeneous dose calculations have been used to prevent worse errors, but, more recently, newer technologies have been implemented in CT reconstruction algorithms that allow for mitigating the effects of metal artifacts. With respect to heavy ion beam therapies, an unexpectedly high biological response to radiation should not fully discredit the modality. A better understanding for the cause should be investigated to see if the challenges can be reduced or even exploited.

2.2.3 Heavy Ions

Heavy ions have a distinct edge in radiotherapy in that they do offer an elevated RBE at therapeutic energy ranges. While protons have a low RBE of 1.1, carbon ion therapy has an RBE on the order of 2.3 [49]. The benefit of high RBE treatments is particularly pronounced for locally advanced tumors; such tumors tend to be partially hypoxic, which makes them radioresistant[1]. One of the advantages of high-RBE treatments is that the damage tends to be more severe in a way that makes them less dependent on proper oxygenation conditions to kill a cancerous cell.

Carbon ions centers are very expensive, and can run on the order of 2 billion USD to construct. However, their benefits of elevated RBE are more pronounced and as such, there is interest in building them even in the United States. Like proton therapy, they have a Bragg peak and are used to treat patients using the SOBP, but also comes with its own set of challenges regarding uncertainties in dose delivery.

2.2.4 Ultrasoft X-rays

Several investigators have published on the effect of lower energy electron irradiation having an elevated RBE [72, 73, 74, 75, 76], which is notable as the conventional thought on photon and electron radiation to be sparsely ionizing in nature, and supposedly, less likely to have a gain in biological effectiveness. It is important to note that these have been observed in experiments and are important effects. In particular, the dose response of cells to

carbon-K shell X-rays of 287 eV and aluminum K-shell X-rays of 1560 eV were demonstrably higher than standard Co-60 photons of 1.17 and 1.33 MeV (find reference). Not only was the RBE elevated, but the cell survival curves bore similarities to the behavior of densely-ionizing alpha particle radiation in that the linear portion of the linear-quadratic behavior was much more pronounced than is typical in photon radiation[72].

The experimental result is significant for at least two reasons: it showed that there had to be a more specific target for radiation action than the cell nucleus itself as was postulated [50], pointing more directly to the DNA itself, and also that individual DNA DSBs are capable of inducing cell death. The earlier approach is more logical to most of the physics community in the way treatment plans are typically produced and prescribed: a gross accumulation of radiation dose leading to cell death is more easily grasped than having to implement a specific nanoscale target model.

The unique property of carbon K-shell X-rays is that they are only capable of photoelectric absorption leading to a photoelectron of 280 eV, itself having a range of only a few nanometers in water. A creative analogy is that carbon K-shell X-rays produce electrons that deposit the equivalent of 280 keV over a few nanometers at the scale of DNA. However, these photoelectrons have such a short range that it is safe to assume they can only induce a single DSB, although it can be fairly complex owing the the concentrated nature of its energy deposition.

Apart from the experimental cell survival experiments of carbon K-shell X-rays, some recent publications are investigating low energy electrons and photons having slightly elevated RBE. This is not too surprising, as these data have been available for some time and are documented in clinical recommendations pertaining to low dose rate (LDR) brachytherapy, particularly in the use of I-125 and Pd-109 seeds [77]. The impact of such studies may not be too high, as clinical data indicates slightly different cell survival data, but clinical prescriptions are based on a long history with the use of these radioisotopes. As such, the clinical prescriptions are more based on the long treatment history of patients specifically

with respect to the radioisotopes in use: the biological response of patients to these have been documented even without specific nanoscale biophysical models. The clinical data does, however, provide much data that can be retrospectively analyzed for consistency. Furthermore, assessment of biological equivalencies or expected response to these radiation types can help bridge the modalities of external beam radiotherapy and brachytherapy more closely together.

2.2.5 Fast neutrons

A special mention of fast neutron therapy should be made as a large premise behind the presented model is to address a specific gap required for safe delivery of new treatment models. Neutron beam therapy, and boron neutron capture therapy (BNCT) had been investigated with some interest in the United States in the 1970's and 1980's [54, 78, 28], but has largely been abandoned due to uncertainties in treatment and substantial secondary effects. Neutron beam therapy was a hopeful modality for treatment of radioresistant cancers. Primary radiation damage from neutron beams are through indirect action; neutrons produced from the acceleration of deuterons with a target of a mean energy on the order of 28 MeV would be focused on a treatment volume and would subsequently interact with hydrogen atoms in the patient, which would be converted to recoil protons depositing energy along the neutron beam. This is similar to external beam photon therapy, except that the recoil protons would have a more pronounced biological effect.

Among the various options for clinical radiotherapy, several authors have made the claim that the highest RBE treatment modality is neutron brachytherapy via Californium-252 [79, 80]. Cf-252 is very heavy radioisotope that decays primarily through alpha decay, but has a relatively high yield of spontaneous fission with a mean neutron energy of 4 MeV and a varying yield of neutrons with each spontaneous fission event. With a half-life of 2.6 years, it is one of the only radioisotopes available that has a low enough activity and a high enough fission yield to be effectively used for neutron brachytherapy. Investigations

of Cf-252 NBT were conducted in the past, but largely abandoned due to the risks and uncertainties involved.

From a radiobiology perspective, though, Cf-252 NBT has a projected RBE of between 4 and 8, far higher even than carbon ion therapy [79]. An additional benefit of NBT is that it is very, very cheap relative to other high-RBE radiotherapy modalities. There, however, is no direct consensus on the efficacy of treatment, and there are uncertainties in terms of the ratio of dose absorption in the patient from different sources. Neutrons absorbed in a patient can lead to subsequent decay of the neutron absorbing nucleus, and the Cf-252 source itself is documented to deposit dose in the form of neutrons as well as gamma decay, with the specific behavior changing with the age of the source itself [81]. The mention of the uncertainties is to show the value of more comprehensive biophysical modeling: an appropriate model can help determine the likelihood of a treatment mode being worthwhile. While there is no guarantee that neutron therapy is the treatment of the future, the possibility exists that it can be highly effective in some cancer types. The biophysical model we present should be able to predict the expected response of irradiated cells for any variation, quantity, and combination of radiation exposure. If it does show such modes are feasible, it can open up opportunities that are not currently under consideration.

2.3 DNA Organization

The organization of DNA is itself a complex topic that is constantly evolving[23]. Even from a perspective of clinical radiotherapy, very few clinical physicists are aware of the extent to which DNA compaction has been investigated and explored over the past several decades. Rather than exhaustively explore the extent of how DNA is organized in the cell nucleus, a rather brief approach to the specific details will be undertaken, with some emphasis on the bioinformatic tools being applied that have yielded the more detailed nuances of subnuclear DNA organization for G1 phase, and also the radiobiological responses that have contributed to the understanding that nanodosimetric modeling is important for

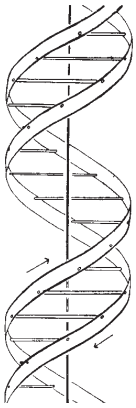
assessing cellular damage by radiation.

Deoxyribonucleic acid, DNA, is the coding behind eukaryotic life and one of the most important discoveries of the twentieth century [82]. With respect to radiobiology, DNA is the consensus primary target for radiation damage [1]. Other cellular components including mitochondria have been observed to be involved in cell death, and other contributors to cell killing such as bystander effects have been noted, but the primary means by which radiation kills cells is through misrejoining of DNA following the formation of DNA double-strand breaks [83]. Much of the evidence behind this was observed at the point of mitosis when chromosomal aberrations from non-homologous end joining (NHEJ) can be seen through a microscope.

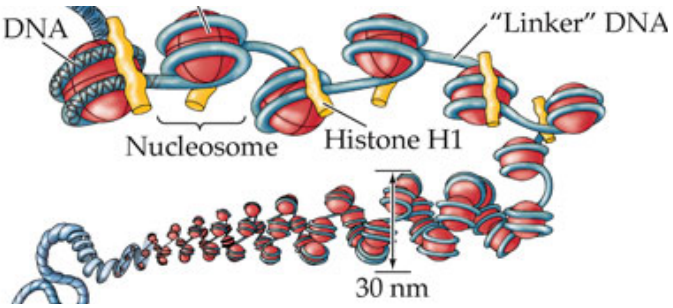
2.3.1 DNA Double Helix, Nucleosomes, Chromatin Fibers, and Chromosome Domains

The double-helix structure of deoxyribonucleic acid (DNA), shown in Figure 2.4 was published in 1953 by James Watson and Francis Crick [82]. Expertise in X-ray crystallography was contributed by Rosalind Franklin, which helped visualize the nanoscale structure of the building block of the eukaryotic genome. While DNA is now understood to be the consensus target of radiation damage, it had already been applied in the treatment of cancers for decades prior [1]. The positive and negative effects of radiation exposure are both attributed to DNA damage and subsequent misrejoining events, including mitotic cell death and carcinogenesis [84].

A sugar-phosphate backbone comprises the 2nm diameter double helix, with matched base pairs in a purine-to-pyrimidine conformation between the two helices [82]. The coding of all genomes is accounted for by the sequence of base pairs, and in a human cell nucleus approximately 6 billion base pairs of DNA encode the entire genome contained within 23 pairs of chromosomes. It has been established that the spacing of 10 base pairs occupies a length of 3.4nm of DNA, meaning that the human genome, if stretched out, would have a length exceeding 1m of DNA.



(a) The DNA double-helix[82]



(b) The Nucleosome and Chromatin Fibers[85]

Figure 2.4: DNA as a double helix, compacted around nucleosomes and into chromatin fibers

The entirety of the DNA sequence is contained within the cell nucleus and is compacted in multiple layers. The initial compaction involves the core octomeric histone proteins around which the DNA double helix is wrapped to form the nucleosome [24]. The DNA wraps around the core histone proteins approximately two times with linker histone H1 at the periphery to form nucleosomes of 10nm diameter as seen in Figure 2.4. This form is also known as euchromatin fiber, and can be condensed further into a 30nm diameter heterochromatin fiber [23]. The specifics of DNA expression is heavily tied to access to DNA, with euchromatin fiber as the more typical form common to transcriptionally active DNA and heterochromatin being more common to less active DNA.

The actual expression of genes is a remarkable and complex field of study, and many investigations aim to illuminate finer details regarding gene expression and their impact on life in the field of epigenetics [24, 86, 87]. Epigenetics is a rich subject and one of great investment in cancer research. The impact of histone proteins on gene expression is enormous, and the importance of the cell's ability to express specific portions of its genome is critically important for its ability to function [24]. To this end, many researchers are continually involved in unpacking even the smallest details of each protein and pathway of the animal kingdom.

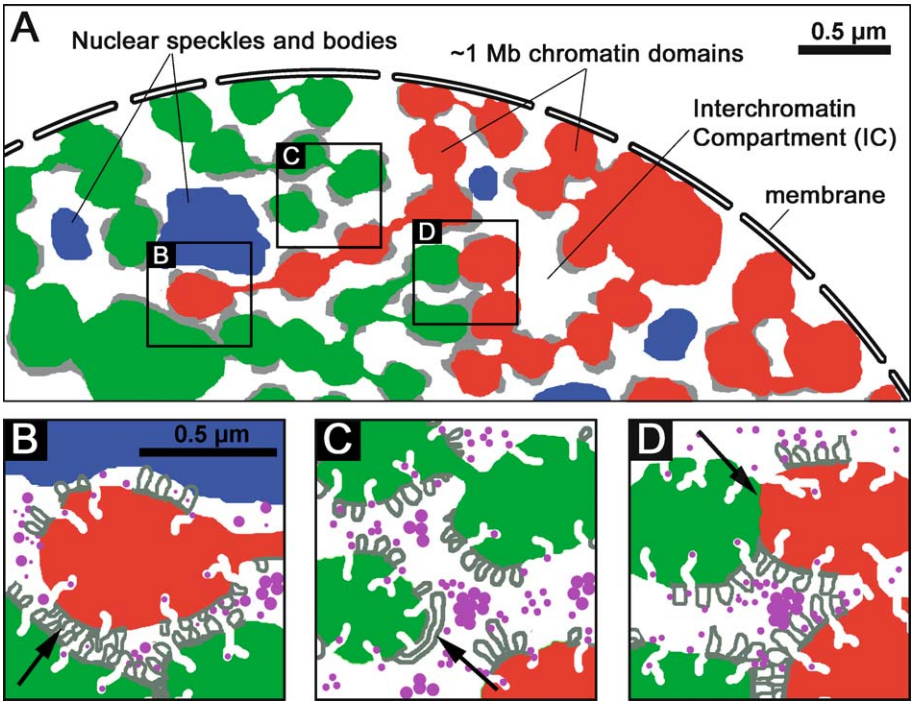


Figure 2.5: DNA compaction into chromatin domains[21]

The most obvious form of DNA organization to the common reader is the fully compacted 1400nm wide chromosome prior to mitosis [1]. However, the majority of a cell's time in the cell cycle is in the G1/interphase state, wherein the DNA is relaxed in such a way as to allow access by various enzymes for normal cellular function. The presented model approximates the G1 cell cycle phase, which itself has a remarkable level of structural organization and compaction beyond that of euchromatin and heterochromatin [21]. Fluorescence imaging in more recent years has shown that chromosomes tend to occupy distinct regions in interphase [89], and Figure 2.5 shows that even within these chromosome territories, there is a further level of compaction into more chromatin-dense chromatin domains (CDs) [21]. These CDs are approximately 400nm in diameter on average and contain on the order of 1 MBP of DNA. Each CD contains chromatin fibers, with euchromatin in CDs that are more transcriptionally active and heterochromatin where there is lower activity. CDs are separated by interchromatin compartments (IC's), which allow for the navigation of enzymes to access different sections of DNA.

2.4 DNA Damage, Chromosome Aberrations, Repair and Misrejoining Mechanisms

Radiobiology is dependent on the common understanding that ionizing radiation damages cells. Damage from ionizing radiation can be caused through direct as well as indirect action: in the case of direct action DNA is interacted with by ionizing radiation, while indirect action involves the interaction of radiation with atoms close to DNA to produce free radicals (OH.) that themselves can diffuse a several nm to interact with DNA. Both methods are able to initiate a single strand break (SSB) in DNA, wherein one side of the sugar phosphate backbone is lysed. It is understood that isolated SSBs are a more minor concern with respect to ionizing radiation damage, as they are easily repaired by the cell's repair enzymes and the preservation of the DNA structure is largely unaltered, making proper repair more likely.

DNA double strand breaks (DSBs) are a much more important damage type in radiobiology, and interactions leading to DSBs are of higher consequence than those causing SSBs. DSBs are produced when both helices of the sugar phosphate backbone are lysed in sufficient proximity that the DNA is broken into two free ends. It is understood that two SSBs within 10 base pairs can cause a DSB, but SSBs farther apart typically do not sufficiently weaken the DNA structure so as to cleave the DNA [90].

DSBs themselves can further be categorized in terms of their own complexity. A simple DSB (sDSB) involves a clean cut of DNA, which is typical of two SSBs leading to a DSB. A complex DSB (cDSB) is formed when DNA is damaged more severely in a manner so as to make repair much more challenging for the cell; this can take the form of many SSBs in close proximity, and possibly include base deletions, as well [74]. DSB severity has been observed as an important factor in radiobiological consequences.

As DNA damage is common in cells, there are many different ways by which cells can repair them. Different repair mechanisms are active to deal with DSBs and SSBs, and some are more prevalent in different cell cycles. The complexity of damage influences

the amount of time it will take to repair a cell: SSBs can be repaired very quickly, while cDSBs can linger for many hours. One of the important elements of DSB repair is that misrepair is an important vulnerability, as improper restitution of broken DNA ends can lead to mutations, chromosome aberrations, and even cell death. This makes the DSB the most important damage type in radiation biology. Furthermore, the most important repair mechanisms are homologous recombination (HR) and non-homologous end joining (NHEJ) [74]. HR involves the recruitment of the diploid sister chromosome in the cell nucleus as a template to copy and repair broken DNA. HR is slower and more accurate, but more common in the DNA replication S phase of the cell cycle and less common in interphase. NHEJ is much less accurate but very fast, and operates by splicing broken ends very quickly: it is also the dominant repair mechanism in interphase [91]. One of the requirements for NHEJ is full decondensation of chromatin fiber around the DSB, which increases the likelihood of misrejoining when multiple DSBs are in spatial proximity to each other.

DSB misrejoining is the most important damage consequence in radiobiology for applications in radiotherapy as well as radiation protection. The pathways leading to chromosome aberrations and cell death are a result of DSB misrejoining. Nonlethal chromosome aberrations may be as simple as chromosomal translocations, insertions or deletions that can lead to the loss of tumor suppressor or housekeeping gene functions or the over expression of oncogenic factors in normal cells, all of which can eventually lead to carcinogenesis[1]. In radiotherapy, the unrepaired and misrejoined DSBs lead to lethal chromosome aberrations (e.g dicentric and accentric ring), which are mainly responsible for the death of cancer cells.

Among other experiments, one noted experiment indicated the S-phase of the cell cycle to be the least radiosensitive owing to the increase in proteins associated with the more accurate homologous recombination repair mechanism. In this experiment, a creative method for halting the cell cycle at specific checkpoints and subsequently inducing radiation dam-

age in the halted state, and subsequently allowing the cells to progress show that cell cycle phases have varied susceptibilities to radiation damage [41]. Generally, it was found that cells are still susceptible to damage during G1 and G2 phase, as the NHEJ pathway is more pronounced, while during M phase, or mitosis, the cell is particularly prone to lethal misrejoining.

A variety of tools have been deployed to test out many of the varying theories, and as pertains to the direct simulation of radiation damage, most of the heavy lifting involves the direct application of such tools to tackle the pathways at work in the cellular response to radiation damage. Advancement of clinical treatment is heavily dependent on understanding which pathways are typically at work for specific cancer types, as this sheds light on the types of adjuvant therapies that are most effective for one patient that would be detrimental to the health of another. Similarly, immunotherapy is one of the very active research topics and for some patient populations, radiation therapy has some effect in helping assist the body in having a favorable targeting response in conjunction with said treatments.

Further advancement of biophysical models may be able to help project acute effects or increased tumor control probability for concurrent chemotherapy and radiotherapy for patients, but at this early stage the added knowledge of DNA compaction is helpful for developing more effective and dynamic models for better understanding the data we have been collecting through experimentation. With the help of fluorescence in situ hybridization (FISH), several research groups have confirmed the organization of DNA within the cell nucleus during the resting G1 phase consists of euchromatin fibers being expressed, heterochromatin fibers in a condensed form, chromatin domains of chromatin dense space, interchromatin compartments of very low DNA concentration allowing the motion of proteins and other complexes through the cell nucleus, and chromosome territories. There is a very dynamic landscape within the cell nucleus itself, which makes for a very challenging environment to fully comprehend or tease out.

2.5 Current Radiobiological Models

One of the challenges in current approaches to radiobiological modeling by physicists is that there are typically a large breadth of correction factors employed to account for deviations of multiple variables[92]. Some of the drawbacks is that these are limited in their application due to their mathematical complexity[93], and these approaches can lose accuracy when too many parameters are adjusted simultaneously. Moreover, these models are frequently based solely on the physics of the charged particle track, rather than assessing the direct biological targets [93]. Oftentimes the fallback for discrepancies is that there is a biological element that needs to be incorporated to make the biophysical model more consistent, and frequently this may take a very mathematical and immensely cumbersome calculation for implementation [93]. On our end, it alludes to the benefit of a structure-based model that can incorporate features of both the charged particle track as well as specific biological targets for the radiation to interact with.

One thing to note, as well, is that several groups have undertaken the task to model the DNA targets very comprehensively for assessing direct DNA damage by charged particles [94, 7]. While this task would have the benefit of incorporating the direct assessment of DNA damage, there are several features of the biological response that is not included in many such simulations, including the dynamic nature of the cell nucleus itself when damaged. In addition, these approaches are heavily, computationally expensive. While the model we present does not include as much detail as these excellent models, we contend that the amount of information we can readily assess from a less detailed cell nucleus model can more expediently produce clinically relevant results through creative statistical methods which will be further expounded upon. The computational efficiency of our biophysical model is such that multiple radiation qualities can be assessed and compared for the production of cell survival curves that are useful in comparing radiation of different types. This can greatly expedite the safe and responsible investigation of clinical radiation

beams when the delivery is of a radiation quality that is not fully understood. Part of the challenge in developing an accurate biophysical model is that it should be able to replicate observed phenomena, regardless of how challenging the underlying principles are. If a simulation can replicate the expected behavior by its inherent function, then it can be applied for the more helpful task of explaining existing scientific theorem pertinent to the field. More importantly, an accurate model can then be applied directly to guiding the scientific community to ask even more challenging and meaningful questions.

Some of the aforementioned models aim to explain the damages by using concentric rings to determine the ratio of dose deposited in the immediate vicinity of the original charged particle track, versus the dose deposited at progressively farther radial distances from the original particle[37]. A few models are based on clustering algorithms to determine how clustered the dose deposition is from all energy depositions, which is more of an assessment of potential damage. In addition, some have begun to include the biological structures common to a cell nucleus [17].

CHAPTER 3

THE MONTE CARLO-BASED SIMULATION METHOD

As mentioned in Chapter 1, the simulation study was divided into four independent parts: (1) development of charged particle track structure data library, (2) simulation of the cell nucleus, (3) production of DNA double strand breaks, and (4) evaluation of cell survival, with the intended flow being modularized so as to allow for minor updates to be implemented in the future. Accordingly, the simulation code is divided into four different parts of which each uses different Monte Carlo algorithms, codes, or coding languages. For example, in part 1 the particle track structure data were generated using the Monte Carlo code Geant4-DNA [6, 30]; in part 2 the chromatin fiber positions were produced using an in-house Monte Carlo code that is run using Geant4 as an engine [18]; and in parts 3 and 4 the production of DNA double strand breaks and the evaluation of cell survival were performed in Matlab for visualization and owing to the friendliness provided by Mathworks. Figure 3.1 is the flowchart showing how the four parts of simulation, as well as the additional details of each part, are put together to produce the final results. Each of the four parts produces its own ”intermediate” results allowing the code to be validated against the results obtained from other methods, be it computational or experimental. The details for each of the four parts are elaborated respectively in Chapters 4-7.

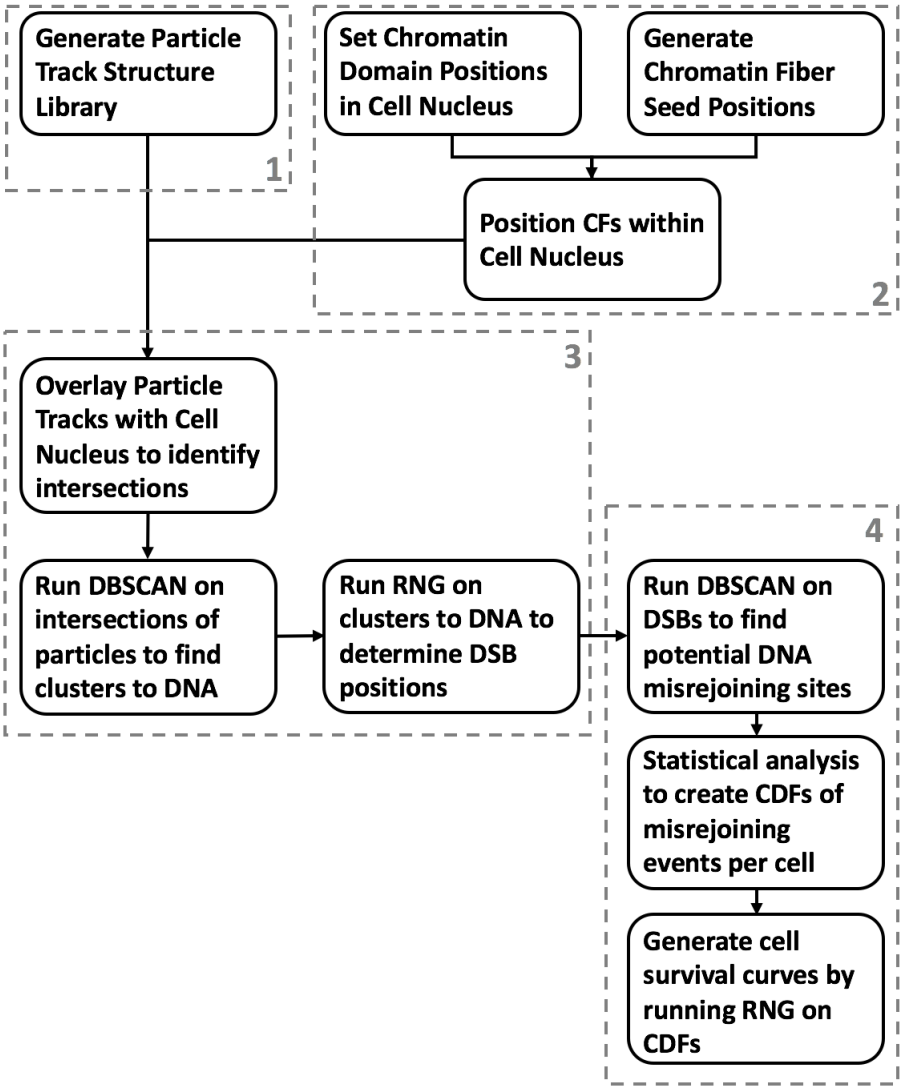


Figure 3.1: The flowchart of the Monte Carlo-based simulation method for calculating radiation-induced DNA damage and cell survival

CHAPTER 4

GENERATION OF PARTICLE TRACK STRUCTURE LIBRARY

Keeping with clinical physics best-practices, our model utilized the Monte Carlo code Geant4-DNA for the production of charged particle tracks [6]. It is important to note the challenges involved in performing simulations at the scale of DNA: Monte Carlo codes are generally understood to be accurate within specific parameters, and very few codes have the capabilities of Geant4-DNA that account for accurate physics processes at a scale relevant for DNA modeling.

There are limitations for performing these calculations, and it is important to be aware that the scope of simulations are currently restricted to certain assumptions. Various codes are optimized for specific purposes, and offer a breadth of particle transport models for various media through which radiation can travel through. Owing to limitations of computing power and realistic expectations of calculation times, many codes assume a certain threshold energy for the secondary electrons along the particle track. That is, electrons with energies below the threshold value are not followed and the energy of the electron is assumed to be deposited locally. [30, 32]. The threshold energy for the secondary electrons assumed by many codes is around 10 keV. While this assumption is good enough for Monte Carlo codes for typical clinical applications, it is not acceptable for our simulation study for two reasons. First, the range of a 10-keV electron in liquid water is as long as 2.5 um, which is much greater than the size of DNA and chromatin fiber. As such, the assumption of local energy deposition of a 10-keV electron would completely miss the most important details as to where the energy is actually deposited at the levels of DNA and chromatin. Second, electrons with energy below 1 keV tend to form clusters of ionization and excitation events within a distance of a few nanometers, which is most relevant to the production of DSB. The Geant4-DNA code is capable of simulating electrons with energies below 10

eV, and therefore, can produced details data regarding the exact location and the amount of energy deposition for each ionization and excitation event along a charged particle track [6, 36].

Liquid water is the current standard medium for nanodosimetric Monte Carlo modeling of charged particle tracks [7, 6]. Several authors have mentioned that their own radiobiological models can likely be made more accurate by incorporating a more tissue-equivalent medium, but this is unlikely to be a reasonable claim. First, liquid water has historically been the assumed medium in clinical radiotherapy by which the human body is modeled [1, 10]. As most cells are comprised of about 70 percent water, this is a relatively safe assumption to make. Furthermore, a tissue-equivalent cross section library is highly unlikely to be made available in the foreseeable near future. In fact, the effort made the currently available cross sections of liquid water for charged particle track simulation has been a monumental task, and the cross sections has only recently be made easily accessible to the general research community via the Geant4-DNA code [6, 30].

For producing charged particle tracks, Geant4-DNA is limited in terms of the energy availability of particles: electrons can only be modeled below 1 MeV, protons below 100 MeV, helium ions below 400 MeV, and heavier charged particles below 1 GeV/amu [6]. The LET produced by electrons, protons, helium and carbon ions were tested and found to be in good agreement with published data, but some divergence was found with the heavier ion tracks such as silicon and iron. This indicates that some of the data extracted from Geant4-DNA may still need further validation in terms of the accuracy.

4.1 Isotropic Source Model for Electrons

One of the hurdles that were encountered for the simulation of charged particle tracks for our biophysical model is in the simple parameter of how to orient and produce the tracks. Initial studies for single track effects simply followed along the central axis of the charged particle, which worked well for heavier ions and with relatively high success

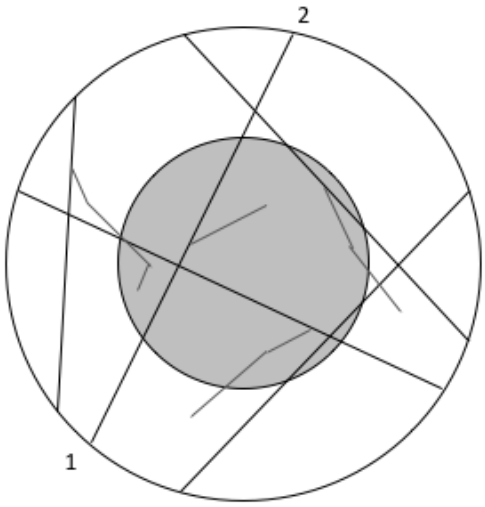


Figure 4.1: Isotropic particle source model for electrons

even for electron tracks [18]. However, since one of our goals was to produce cell survival curves which are heavily dependent on interactions of damage from neighboring particle tracks (especially in the case of sparsely ionizing electrons), this precluded the single-track approach and a more complex means by which the production of randomly oriented charged particles would be required.

For electron irradiation modeling, an isotropic source approximation was applied. As shown in Figure 4.1, the two points (i.e. 1 and 2) corresponding to a straight line intersecting the spherical cell boundary were selected at random and each electron was initiated at one point (i.e. Point 1) and given an initial direction toward the second point (i.e. Point 2). This process was repeated for as many electrons as needed until the specified dose of radiation was delivered for the entire cell nucleus.

The premise behind the modeling approach used for sparsely ionizing electron radiation was that clinical photon and electron beams can be assumed to behave isotropically owing to the sparse nature of their energy deposition patterns and the effect of large an-

gles of deflection of primary electron interactions. In most cases of clinical interest, this assumption holds true with certain exceptions where photon energies are extremely low as is the case of carbon and aluminum K-shell X-rays [95, 73]. Clinical treatment beams of external radiation are delivered in a source-to-target type of orientation, with all of the particles emanating from a single source and delivered in the direction of the target of radiation [62]. This method is also true for in vitro and in vivo experimentation for most cases: several products are in common use specifically for cell irradiation and small animal studies, and they can make use of low energy X-ray sources in addition to higher energy radionuclide sources of Cs-137 and Co-60, but in both of these cases they are typically set up with the radiation source at a fixed distance and irradiating towards the target, although they may make use of different beam geometries. Moreover, even brachytherapy and nuclear medicine procedures make use of sources energetic enough that the isotropic approximation should still be relevant.

4.2 Beam Line Source Model for Heavy Charged Particles

Heavy charged particles (or ions) tend not to diverge as much and have a more densely ionizing and clustered energy deposition pattern, especially as the particles lose energy and their LET increases [96, 49]. Owing to this property, the degree to which interactions would be likely to interact with each other is heavily dependent on the proximity of charged particle tracks and the effect of source origin and particle travel has some relevance in modeling. For clinical carbon and proton therapy beams, in addition to microbeam irradiations for cell survival studies, the ions are accelerated and delivered from a source position towards a target. As such, particles are transported along a beam line towards a target with small angular deviations relative to the central axis of the beam as seen in Figure 4.2.

Accordingly, heavy charged particles were modeled assuming they are transported in the same direction with small angular deflections randomly assigned to the particles with initial positions determined from a planar source using a uniform random number gener-

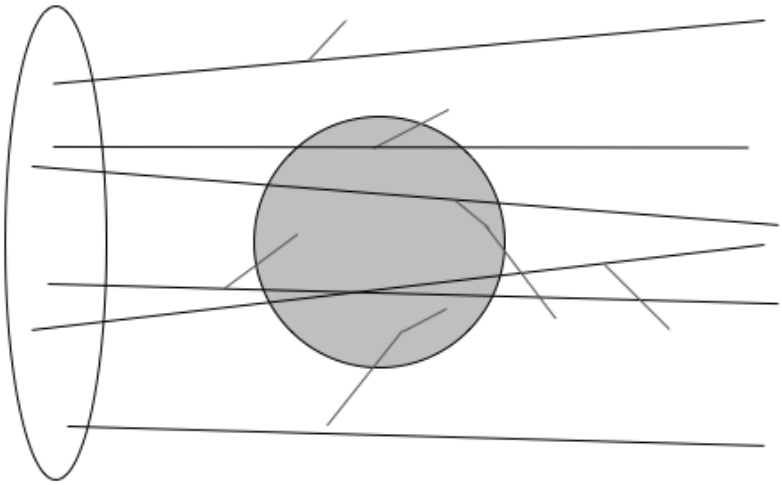


Figure 4.2: Beam-line particle source model for heavy charged particles

ator. In addition to providing some means of angular deflection, the planar disk source (shown in Figure 4.2) was modeled as being larger than the diameter of the cell nucleus to account for in-scattered delta rays from adjacent heavy charged particles to compensate for delta rays escaping from the cell nuclear volume.

It is important to note that while the application of the model follows the behavior of charged particle transport as highlighted in this section, the model itself is not limited to these methods or even Geant4 for production and simulation of charged particle tracks. Any Monte Carlo code can be used for simulation of charged particles with any particular beam arrangement.

4.3 Compilation and validation of track structure data library

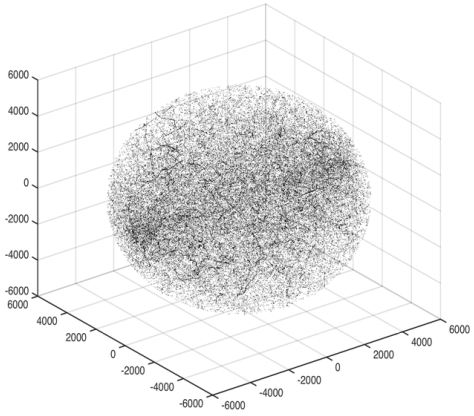
Libraries were compiled using Geant4-DNA based on specified parameters for the particular charged particle type. The specifications for each set of data was: charged particle type, charged particle initial energy, and dose to the cell nucleus. As mentioned in Sections 4.1

and 4.2, the beam-line source model was used for simulation of heavy charged particles, and the isotropic source model was used for electrons. For each set of data, a total of 30 independent irradiation schemes were produced. For example, a total of 30 unique sets of 0.5 Gy dose, 100 MeV proton irradiation were stored into the library, as were 30 unique sets of 1.0 Gy dose, 10 MeV proton irradiation. Figure 4.3 is a graphic display of two different particle type data sets of 1 Gy of dose: one for 1-MeV electrons and the other for 17-MeV/amu carbon ions. The difference in spatial distribution of energy deposition events between the two particle types can be clearly seen.

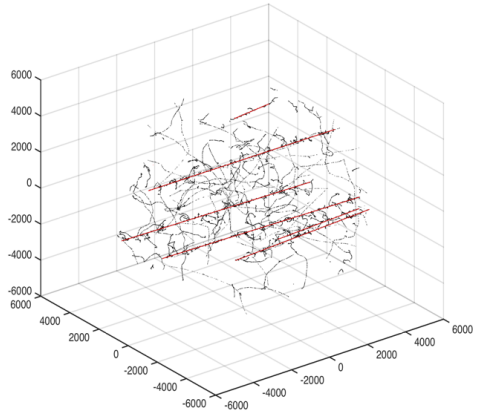
Settings within Geant4-DNA were performed for appropriate simulation of charged particles, and physics lists and particle generators within Geant4-DNA were confirmed for proper function. Additionally, as Geant4-DNA makes use of discrete history instead of condensed history for Monte Carlo simulation, output files from Geant4-DNA examples involving all charged particle interactions are inherently unwieldy as they include all elastic scattering interactions which tend to compose the majority of electron interactions Modifications to the code were performed such that only excitation and ionization events capable of causing DNA damage were recorded. This threshold for interactions is 7 eV, and any events below 7 eV were not stored.

4.3.1 Data Output

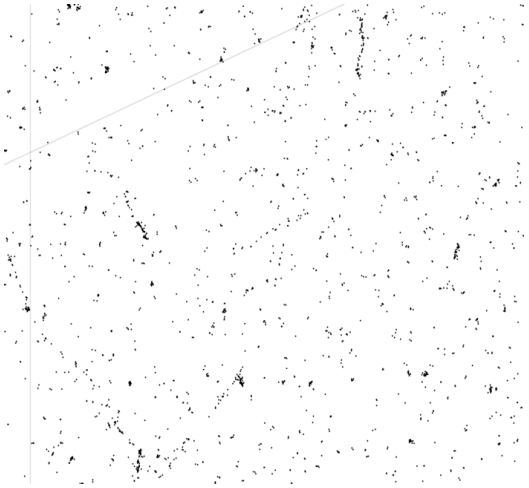
Data output selection from Geant4-DNA was a major task to accommodate, as several parameters of little interest for many researchers were critically important to the type of nanodosimetric analysis we deemed to be essential. One of the powerful elements of Geant4 is how customizable the software is: it is a monumental simulation tool that is extremely unwieldy, but the lack of a simple interface enables the user to tab even obscure data that is normally considered inconsequential to the end user. This was crucial for the simulation of the cell nucleus model in tandem with the mean free path approach applied to the positioning of chromatin fibers.



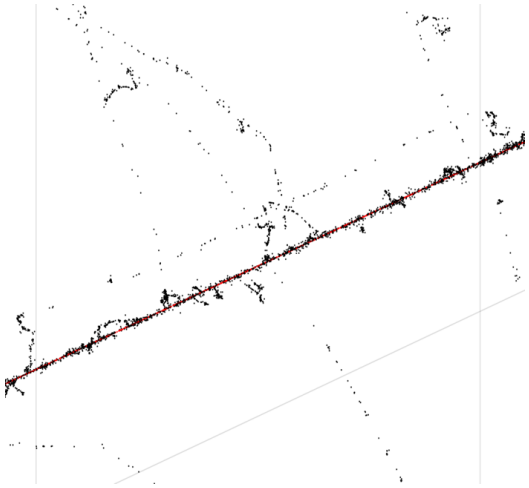
(a) Isotropic 1-MeV electron tracks



(b) Beam-line 17 MeV/amu carbon ions



(c) 1-MeV Electron tracks, zoomed in



(d) 17 MeV/amu carbon ions, zoomed in

Figure 4.3: Visual comparison of the spatial distribution of energy deposition events of 1 Gy dose of two radiation types: 1-MeV electron ans 17-MeV/u carbon ion

The output data for charged particle tracks from Geant4-DNA were specific for this simulation study. Data were stored in comma-separated-value text files for every energy deposition event above the 7 eV threshold. The cartesian coordinates, total energy deposition, and particle kinetic energy are stored for each event. The individual step length for each event in addition to the cumulative step length are also tagged for each event, with the individual step length used for positioning of the events and the cumulative values being used to confirm proper tracking for significant energy depositions between events. This is confirmed externally in Matlab to verify the individual step lengths for events cumulatively match the values output in Geant4, and is necessary for the positioning of chromatin fibers in the cell nucleus model generation. Additional modifications were made to account for insufficient significant digits for event numbers and were later accounted for in post-processing.

The event identifier, particle identifier, and parent particle identifier are also output for each event. Essentially, each primary charged particle simulated in the Geant4 charged particle generation itself is a unique event, with the particle identifier of 0 denoting the originating charged particle. Subsequent delta rays would identify their parent particle from which it was produced, while also identifying which originating primary charged particle it is associated with. Identification of individual, unique charged particles allows for identification of how DNA DSBs are generated in our cell nucleus model: essentially we are looking into the α and β coefficients of the linear quadratic equation. This also allows for discriminating finer details into how DSB misrejoining events are occurring based on which particles and originating events are involved in producing the DSBs.

4.3.2 Dose Consistency

As a validation method for Geant4, the LET values for various heavy charged particle types obtained from Geant4 were compared with the expected values obtained from the SRIM code [97]. As shown in Table 4.1, these values do not match exactly but are in reason-

Table 4.1: Comparison of the LET values obtained from SRIM and Geant4 for various particle types and the corresponding value of number of tracks per Gy.

	SRIM LET (keV/um)	Geant4 LET (keV/um)	Tracks/Gy
1 MeV electron	0.20		2425
28 keV photons			1845
aluminum K-shell X-rays			5112
carbon K-shell X-rays			16285
Protons, 100 MeV	0.74	0.72	1012
Protons, 10 MeV	4.7	4.7	130
Protons, 3 MeV	12.1	12	48.7
Protons, 1 MeV	24	36	15.9
Helium, 400 MeV	2.9	2.4	300
Helium, 20 MeV	32.6	29.7	19.9
Helium, 5 MeV	91.2	105	5.4
Carbon, 290 MeV/amu	12.7	10.3	61.8
Carbon, 185 MeV/amu	16.8	13.5	45.9
Carbon, 67 MeV/amu	35.4	33.0	21.3
Carbon, 17 MeV/amu	111	99.4	6

able agreement. Since at this stage we have chosen to perform our analysis through the libraries within Geant4-DNA, we opted to instead match the quantity of charged particles with the total energy loss simulated in Geant4 for establishing libraries of total dose for each charged particle. Hence, while the total LET varies in Geant4-DNA relative to referenced publications, the total dose deposition is still kept consistent by modifying the tracks per Gy called out by our model. As a reference, Table 4.1 also shows the number of tracks per Gy for the various particle types.

It should also be noted that the total energy lost by primary charged particles is consistently higher than the sum of the total energy deposition within cells from our output files, which is expected given that a lower energy threshold of 7 eV is used to filter out events incapable of producing DNA damage. Hence, energy depositions below 7 eV are ignored in our simulations for all particle types. For heavy ions including protons, the total energy lost by the primary particle can be observed by noting the initial and final kinetic energies of the particle. It was consistently noted that the total of recorded energy depositions was consistently at 85% of the change in kinetic energy of the heavier ions in our simulation. Hence, for dose confirmation purposes, we verified that the summation of individual energy deposition events in the cell nucleus was 85% of the dose deposited to the cell.

For electron and photon radiation, the 85% ratio of energy deposition to dose was assumed to adjust the quantity of particle tracks. Unlike heavy charged particles, electrons are capable of considerable deflection from their initial direction and are also capable of transferring very high proportions of their kinetic energy in hard collisions, and as such it is not possible to use the primary charged particle kinetic energy loss for determining their dose deposition. Since the total sum of energy depositions from charged particles underestimates the total dose deposition within the cell nucleus, a correction must be applied to adjust the apparent dose deposition to the actual dose deposition.

CHAPTER 5
SIMULATION OF AN INTERPHASE CELL NUCLEUS

After compilation of the particle track structure database, the next step of the effort is to simulate the cell nucleus consisting of the coordinates (i.e. positions) of enough detailed "subnuclear components" (i.e. chromatin fiber, chromatin doain, and interchromatin compartment as described in Section 2.3) from which the particle track structure data can be used to produce DNA damage. The primary basis for subnuclear structures and regions have been largely assumed from publications based on FISH imaging and electron microscopy [21, 22, 98].

5.1 Cell nucleus components, CDs, IC, CFs, etc.

The primary structures of interest in the cell nucleus model is in the DNA compaction on the order of chromatin fibers (CFs), chromatin domains (CDs), and the interchromatin compartment (IC). While in a real cell nucleus CF is a continuous piece of strand, in our model the CF is broken into many discrete pieces of "rod" so that they can physically fit into the discrete CDs. Figure 5.1 shows the general layout of the CDs and CF rods within the modeled cell nucleus. As shown, the cell nucleus is divided into two regions, the region of CDs and the region of IC. This depiction of a cell nucleus is consistent with that described by Albiez et al. [21]. Within the model coding, it is noted that the primary difference between the CDs and IC is that the CF rod density is significantly higher within CDs than in the IC, and this is shown in Figure 5.1 in the different proportions of CF rods in both regions. Details into the concentration and quantites of CDs and CF rods are shown in Table 5.1.

One assumption of our model in using such an approach is that the majority of energy deposition by ionizing radiation causes no damage to the cell. Figure 5.1 shows this in that

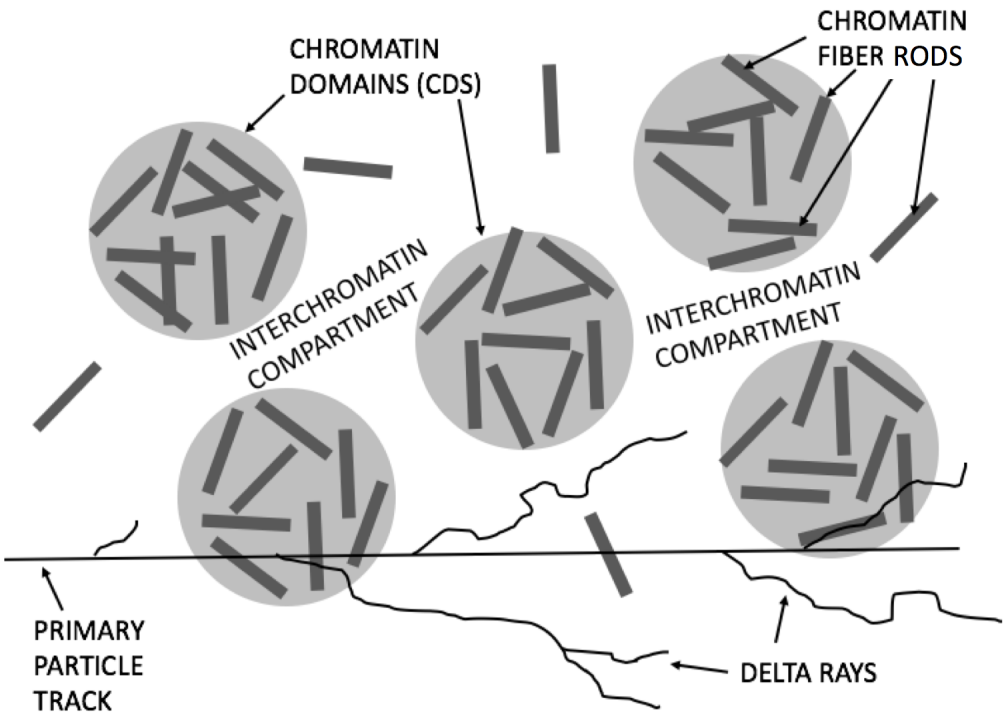


Figure 5.1: Layout of modeled DNA structures with ”a” charged particle track

Table 5.1: Cell Nucleus Components

	Value	Units
DNA content	6	giga-base pairs
DNA in a CF rod	18	kilobase pairs
CF rods in cell	333333	qty
volume of cell	696.9	um^3
CDs per cell	6000	qty
volume of CDs	201.1	um^3
% volume CDs	29	%
% volume IC	71	%

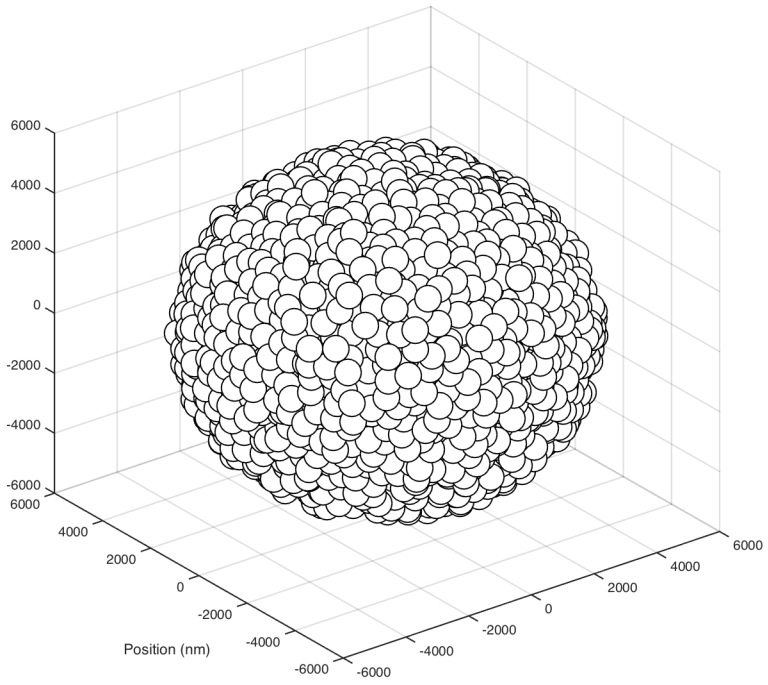


Figure 5.2: Non-intersecting chromatin domains (CDs) within the cell nucleus

the majority of the energy deposition events along a particle track are in the CD and the IC regions that are devoid of CF rods. Only a very small subset of energy deposition along a particle track intersects CF rods, and of the energy deposited to CF rods an even smaller proportion will lead to DNA DSBs.

5.1.1 Sphere packing for CDs and IC

The first stage to generating the DNA structures and regions within the cell nucleus is to define the different regions of chromatin domains (CDs) and the interchromatin compartment (IC). The code is assembled such that the general parameters of a $11\,\mu\mathrm{m}$ diameter spherical cell nucleus is kept with approximately 6000 CDs of 400 nm diameter containing 1 Mbp of DNA each [22] as can be seen in Figure 5.2. None of the CDs are intersecting with each other, and the IC is simply defined as the region of the cell nucleus devoid of CDs. The code is run to compile a library of 150 unique CD/IC matchings.

The stochastic approach to the model for dividing the cell nucleus between an IC and a multitude of CDs is a basic non-intersecting sphere packing algorithm. The code has been written using Matlab, and is a simple application of the Monte Carlo method. In effect, a cubic volume of lengths of $11\,\mu\mathrm{m}$ is defined, and a uniform random number generator populates the cartesian coordinates of points within the cube. For seeding 6000 CDs within a $11\,\mu\mathrm{m}$ diameter sphere, it is determined that the total number of points to be generated within an $11\,\mu\mathrm{m}$ cube would be approximately 11460. A point is stochastically placed, with each new point verified to be no closer than a total distance of 400 nm to any other previously generated point within the cubic volume. If a point is found to be within 400 nm of any previous point, it would be intersecting another CD and is thus discarded, and the process is repeated.

Once all of the points are positioned within the cubic volume, a distance exclusion is performed on the generated points such that only the points that are within a $5.3\,\mu\mathrm{m}$ radial distance from the center of the cube are kept. In this manner, only the CDs that can be fully contained within an $11\,\mu\mathrm{m}$ diameter cell nucleus are kept. The center points of each CD are then output into a data file, and the process is repeated until a total of 150 unique CD/IC template patterns have been generated.

CD volumetric data is not stored, as the algorithm runs more quickly through a k-nearest neighbors search within Matlab based on the center of the CD. As such, the code runs with the CDs being defined as a 200 nm distance from the CD point. Moreover, the individual CD identification is handled based on the order of points: the first point of the output file is CD #1, the second point is CD #2, and so on.

5.1.2 Cylindrical CF rods, quantity and concentration

The total quantity of DNA within a human cell nucleus is 6 Gbp of DNA, with these mostly contained within specific chromatin dense regions of the CDs, each containing an average of approximately 1 Mbp of DNA [22]. CF rods are defined within the model as 30nm

diameter cylinders of 150 nm length. This corresponds to approximately 18 kilo-base pairs (kbp) of DNA [17]. The ratio of these numbers would indicate that, if all of the DNA were contained within CDs, then each CD would contain 56 CF rods. Some of the CF rods would invariably be located within the IC, for basic reasons such as the linkage between CDs from one to another, as well for access by various activities including transcription and DNA repair mechanisms.

The literature indicates that at least a ten fold increase in compaction of CDs over the IC occurs [21]. Furthermore, it is understood that the 6000 CDs are themselves interconnected into their own respective chromosome territories [22]. For our model, that would indicate these CDs are likely connected by CF rods traversing the IC with a minimum length at least equal to the distance between CD boundaries. This would indicate that around 5950 30nm diameter CF rods with a typical minimum length of 120nm. In practice, this is approximately 5% of the proportion of CF rods in the cell nucleus. While this is an important quantity for a very detailed analysis, as further sections will detail, our current assessment of the model is somewhat neutral to large variations in this quantity.

As such, we assumed the compaction rate at a ratio of 20:1 between the IC and CDs. Given that CDs in our model occupy 30 percent of the cell nucleus, and the IC occupies 70 percent of the cell nucleus, the relative compaction ratio would be place about 90 percent of the chromatin fibers within CDs and 10 percent in the IC.

Placement of CF rods in the cell nucleus was performed without concern for overlap within the cell nucleus. The code itself does not allow for double-counting of individual energy depositions, so overlap of CFs with one another would lead to a slightly lower relative volume of DNA within the cell nucleus model than projected, but future sections will demonstrate that this is negligible. Since the CF rods occupy about 5 % of the total volume of the cell nucleus, there would be a very low occurrence of overlap even though it is not accounted for. Moreover, the intersection algorithm, which will be detailed in the following chapter, has been developed to prevent double-counting of damage by ionizing

radiation where the CF rods may be overlapping.

As a whole, the model was approximated as being comprised entirely of heterochromatin fibers and we chose not to incorporate the more relaxed euchromatin fibers at this stage [98]. Part of this decision is intentional as the initial stage of assessing likely DNA damage still holds relevance without a comprehensive modeling of the cell nucleus. Moreover, there are numerous uncertainties for correlating likely DNA damage differences between heterochromatin fibers and euchromatin fibers within the scope of charged particle track energy deposition patterns that could possibly destabilize the model without a more concerted effort. As such, one intention for the future development of the model involves the incorporation of euchromatin modeling, which would be simple given the nature of the mean-free-path approach to the model. However, as the scope of research for the model is itself inherently ambitious, it was determined that a step-wise approach for starting with heterochromatin is sufficiently appropriate. In addition, the current state of the art indicates that our model is novel and has high potential for progressing the field of radiobiology in its present implementation.

5.2 Monte Carlo placement of CF rods

The next subsections detail specifics into the placement of CF rods, and they involve visualization and calculations detailing the finer nuances of our model. It is important to keep in mind that the sole purpose of this section's work is to place all CF rods in the cell nucleus in a consistent, stochastic manner. All of the mathematical rationale of this section simply explains that our model is placing CF rods in such a way as to replicate their placement appropriately.

Figure 5.3 demonstrates the key elements of the Monte Carlo placement of CF rods in our model. A grid of parallel lines is used for mapping out the points at which CF rods are placed. Distinctions in the cell nucleus are made between the CDs and the IC, as the CF rod density is higher in the CDs than in the IC. The CF rods are subsequently placed in the

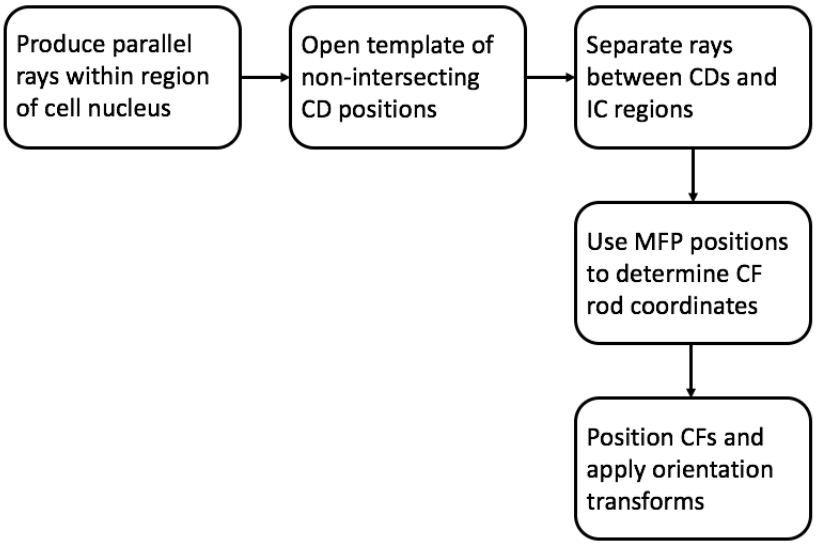


Figure 5.3: The flowchart for CF rod positioning

3-D space of the cell nucleus. The CF rods are then oriented with angular transformations applied.

5.2.1 MFP approach and intersection cross sections of CF rods

A mean-free-path (MFP) approach allows flexibility in placing CF rods randomly within the cell nucleus. Calculation of the MFP for placement of CF rods varies from the previous work [18], in that the basis for calculation is taken from a fixed geometry based on a dense, uniform grid of parallel lines which is detailed further in section 5.2.2. Prior to delving into finer details, a brief introduction is needed for what is meant by the MFP for placement of CF rods.

The MFP approach is taken from principles of Monte Carlo radiation transport calculations. The MFP is the average distance between interactions, which can be used in conjunction with random number generation (RNG) and probability distribution functions (PDFs) to determine how far apart interactions occur and what types among a variety of interactions will take place. This approach was inspired from the ”random walk” procedure used in the Monte Carlo method to randomly determine the distance to the next collision point in neutron/photon transport calculations [4]. The distance to the next collision site is

determined as:

$$S = -\lambda \ln(\mathcal{R}) \tag{5.1}$$

where $\mathcal{R}$ is a random number having a value between 0 and 1, and λ is the MFP of the neutron in the medium. It has been shown that λ is simply the inverse of the macroscopic cross section (Σ) of the particle (a neutron or photon) in the medium, and Σ can be calculated as:

$$\Sigma = N\sigma \tag{5.2}$$

where N is the atomic number density of the medium and it has the unit of atoms cm^{-3}, and σ is the microscopic cross section of each atom for the interaction and it has the unit of cm^2 $atom^{-1}$. Accordingly, Σ and λ have the unit of cm^{-1} and cm, respectively.

Adaptation of the MFP approach to the placement of CF rods is based on the geometric cross section of the CF rods. There is a certain probability that a line traversing the cell nucleus will intersect with multiple CF rods: thus, the average distance between the CF rods along the line is the MFP. To apply the above procedure to a charged particle track intersecting a CF rod in a cell nucleus, one first calculate the geometric cross section of a CF rod. Given a cylindrical diameter of 30nm and a length of 150nm, a CF rod's geometric cross section is simply:

$$\sigma = (0.03\mu m)(0.15\mu m) = 0.0045\mu m^2 CF^{-1} \tag{5.3}$$

In the example of a line traversing a single CD of 400nm diameter with an average of 53 CF rods uniformly and randomly distributed within the CD, the number density of CF rods in the CD can be calculated as:

$$N = \frac{53CFs}{\frac{4}{3}\pi(0.2\mu m)^3} = \frac{53\ CFs}{0.03351\ \mu m^3} \approx 1582\ CFs\ \mu m^{-3}\ . \tag{5.4}$$

Σ and λ can then be calculated, respectively, as:

$$\Sigma = N\sigma = (1582\ CFs\ \mu m^{-3})(0.0045\ \mu m^2\ CF^{-1}) = 7.12\mu m^{-1} \tag{5.5}$$

and

$$\lambda = \frac{1}{\Sigma} = \frac{1}{7.52\mu m^{-1}} = 0.141\ \mu m = 141\ nm \tag{5.6}$$

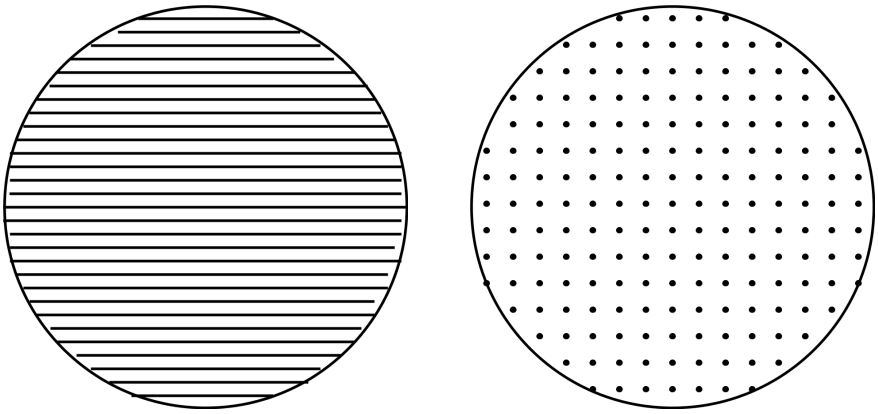
One can then pick a random number $\mathcal{R}$, and apply equation 5.6 to determine S, the distance for the particle to intersect the next CF.

It is important to note that the cross sections applied to CF rods in our model are substantially smaller than the approach along a single line as described above. The reasons will be clarified in the next sections, but are largely due to overlap and to prevent placing excess CF rods in the cell nucleus as . However, the interaction cross sections are scaled such that the actual density and quantity of CF rods matches the expectation values for a cell nucleus.

5.2.2 Grid of parallel lines for determining positions of CF rod placement

The grid of parallel lines is used for transforming the MFP positions of CF rods, a 1-D calculation, into the 3-D space of a cell nucleus. The MFP approach as described in Section 5.2.1 is a quick calculation that outputs only the scalar distances between CF rods, hence it is linear and 1-D. It is obvious that some means for transforming the 1-D positions into three dimensions is necessary.

The grid of parallel lines was determined to be a simple and logical method to transform the 1-D positions into the 3-D space of the cell nucleus for our MFP approach. The MFP approach inherently operates along a single, straight line. Instead of using a single line, a series of lines in a fixed-grid pattern, covering the full cross section of the cell nu-



(a) Grid of parallel lines traversing cell nucleus, side view

(b) Grid of parallel traversing cell nucleus, cross sectional view

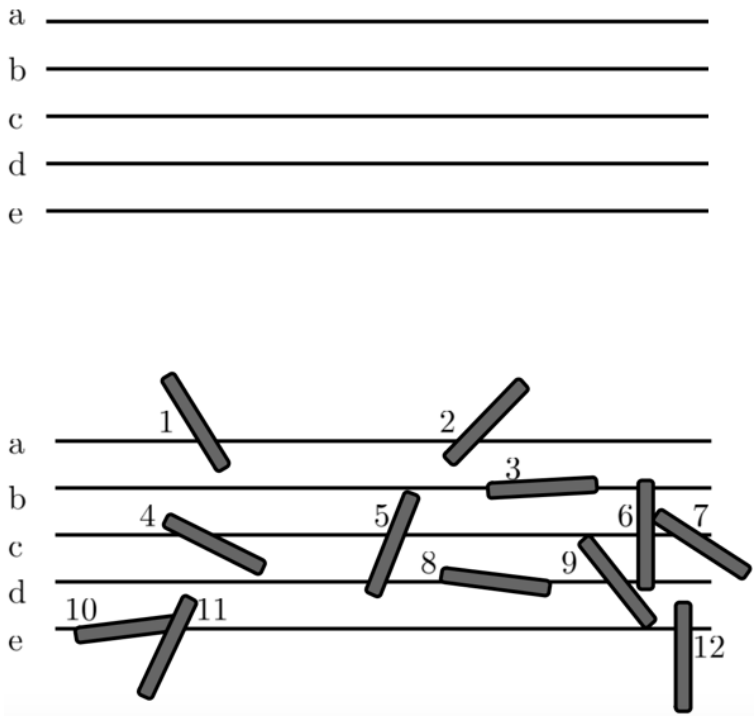
Figure 5.4: Grid of parallel lines used for application of MFP approach to stochastic placement of CF rods in cell nucleus

cleus, is used as seen in Figure 5.4. The fixed grid has a 20nm spacing between lines in horizontal and vertical directions, which gives a possibility for 30nm diameter CF rods to be positioned at any point in the cell nucleus.

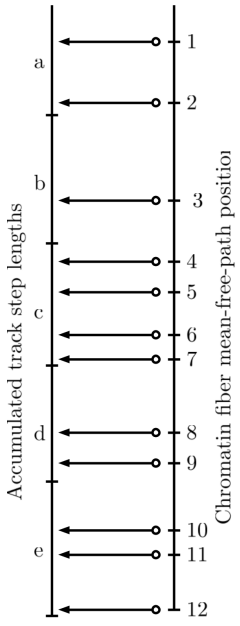
Placement of CF rods along the grid of parallel lines are shown in Figure 5.5. The grid of parallel lines are organized sequentially to match the 1-D MFP placements of CF rods as in Section 5.2.1. This process can be likened to an end-to-end alignment of line segments into a single, long line, or even similar to stretching out cellular DNA from its 3D structural organization into a single double helix. The MFP seed positions for each CF rod are then matched to a point along the grid of parallel lines; thus, each CF rod is planted at the 3-D spatial coordinates of the cell nucleus.

5.2.3 Separation of tracks into CDs and IC

Our model makes use of a cell nucleus of different CF rod concentrations in the CDs and versus the IC. As such, a distinction has to be made between CF rod density in the CDs and the IC. For this purpose different MFP values are used for the different CD and IC regions, and the grid of parallel lines is similarly divided into sections occurring in the CDs and IC as shown in Figure 5.6.

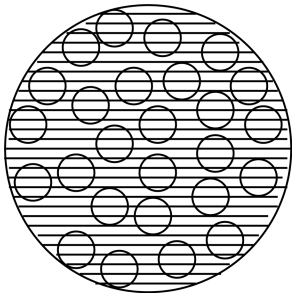


(a) Linear grid view for MFP placement of CF rods

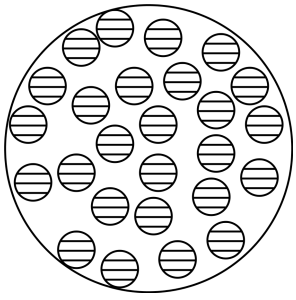


(b) Ladder method for determining MFP placement of CF rods

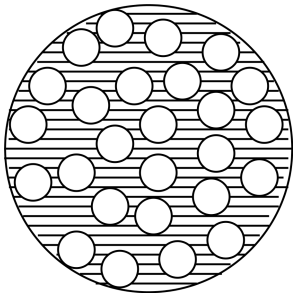
Figure 5.5: MFP placement of CF rods along fixed grids with stochastic seeding



(a) Grid of parallel lines traversing CDs and IC



(b) Grid of parallel lines intersecting only CDs



(c) Grid of parallel lines intersecting IC only

Figure 5.6: Separation of MFP segments by CDs and IC for use of different CF rod densities

In practice this method is very simple to implement: points along the grid of parallel lines are processed through the Matlab knnsearch function against the CD center positions as generated in 5.1.1 to find the distance to the nearest CD center. Any line segment that falls within a 200nm radius from a CD is attached to that specific CD, and any line segment that is more than 200nm from a CD is tagged as occurring within the IC.

The grid lines within CDs are separated from grid lines within the IC at this stage such that the positioning of CFs can be handled with different MFP values for the different densities of DNA within these two regions. The length of grid lines are cumulatively summed and attached to each grid line position for the CD positions and IC positions, respectively. Placement of the CF rods is performed in the same manner described in Section 5.2.2 and depicted in Figure 5.5. In addition, original data regarding the specific CD from which CF rods originate are tagged. Specifics into the parallel lines used and MFPs as detailed in Section 5.2.1 are shown in Table 5.2.

Table 5.2: Cell Nucleus Components

	Value	Units
Mean chord length, sphere	7333	nm
cross sectional area of cell	95033098	nm^2
area per chord	400	nm^2
number of chords	237583	qty
total step length of lines	1742273	um
mfp, uniform assumption	5.227	um
step length CDs	502654	um
step length IC	1239619	um
90% CFs in CD	300000	qty
10% CFs in IC	33333	qty
CD MFP	1.676	um
IC MFP	37.19	um

5.2.4 Orientation of CF rods

Once the mean free path is determined for the intersection of a particle track with the cell model, each CF rod must be positioned and aligned relative to its seeding position along the grid of parallel lines. The major components for alignment are the position offset of the central axis of the CF rod (i.e., the central line of the cylinder's cross section), the rotational of the CF rod, and the final location of the two end points of the CF rod.

CF rod position offset from placement point along a line

The explanation for placement of the CF rod is taken from earlier work in the master's thesis [18]. The placement of CF rods along the grid of parallel lines requires a positional offset from the intersection point for the purpose of random placement: without an offset, all CF rods would be predictably aligned at the center of the grid positions. A simple solution based on the MFP entry position of a line with the CF rod is to isotropically select the direction for the CF rod in the forward direction. For this, the polar angle θ and the azimuthal angle ϕ must be determined stochastically. The polar angle's probability distribution function is given by:

$$f(\theta)d\theta = \frac{dA}{A} = \frac{r^2 d\Omega}{A} = \frac{2\pi r^2 \sin\theta d\theta}{2\pi r^2} = \sin\theta d\theta \tag{5.7}$$

where dΩ is the differential solid angle:

$$d\Omega = 2\pi \sin\theta d\theta \tag{5.8}$$

Integrating the probability distribution function of the polar angle θ yields:

$$F(\theta) = \int_0^\theta f(\theta)d\theta = \int_0^\theta \sin\theta d\theta = 1 - \cos\theta = \mathcal{R} \tag{5.9}$$

Where R is a uniform random number with value from 0 $\sim$1. This yields:

$$\theta = \arccos(1 - \mathcal{R}) = \arccos(\mathcal{R}) \tag{5.10}$$

The azimuthal angle θ has a probability distribution function of:

$$f(\phi)d\phi = \frac{d\phi}{2\pi} \tag{5.11}$$

Integrating the probability distribution function of the azimuthal angle yields:

$$F(\phi) = \int_0^\phi f(\phi)d\phi = \int_0^\phi \frac{d\phi}{2\pi} = \mathcal{R} \tag{5.12}$$

Simplifying this yields:

$$\phi = 2\pi\mathcal{R} \tag{5.13}$$

where R is a uniform random number with value from 0 to 1. The unit vector to the CF rod's center can be calculated in cartesian coordinates as:

$$\hat{\Omega} = u\hat{i} + v\hat{j} + w\hat{k} \tag{5.14}$$

where the unit vectors u, v, and w are calculated by:

$$u = \cos\theta \tag{5.15}$$

$$v = \cos\phi\sin\theta \tag{5.16}$$

$$w = \sin\phi\sin\theta \tag{5.17}$$

The center of the CF rods were then calculated by multiplying the unit vectors by 15nm, the radius of the CF rods:

$$x = r \cdot u = r\cos\theta \tag{5.18}$$

$$y = r \cdot v = r\cos\phi\sin\theta \tag{5.19}$$

$$z = r \cdot w = r\sin\phi\sin\theta \tag{5.20}$$

where r is the radius of the CF rod.

CF rod angular rotation

Figure 5.7 shows the geometric orientation of a CF rod with respect to the particle track intersecting it. As shown, the orientation of the CF rod is determined based on two direction vectors, $\vec{\Omega}$ and $\vec{\Omega}'$. The vector $\vec{\Omega}$ is determined by two points, the track entry point, (x_1, y_1, z_1), and the center point of the CF rod, (x_2, y_2, z_2). Its components, u, v, and w, can be calculated by using equations 5.7 - 5.17, with a radius of 15 nm for the CF rod. After u, v , and w have been determined, one can apply the equations below to obtain u', v' , and w' for $\vec{\Omega}'$ [99]:

$$u' = u\cos\theta + (uw\cos\phi - v\sin\phi)\left(\frac{\sin\theta}{\sqrt{1-w^2}}\right) \tag{5.21}$$

$$v' = v\cos\theta + (vw\cos\phi + u\sin\phi)\left(\frac{\sin\theta}{\sqrt{1-w^2}}\right) \tag{5.22}$$

$$w' = w\cos\theta - \cos\phi\sqrt{1-\cos^2\theta}\sqrt{1-w^2} \tag{5.23}$$

Since $\vec{\Omega}$ and $\vec{\Omega}'$ are perpendicular to each other, this makes $\theta = 90°$. Equations 5.24 - 5.26 are then reduced to:

$$u' = \frac{uw\cos\phi - v\sin\phi}{\sqrt{1-w^2}} \tag{5.24}$$

$$v' = \frac{vw\cos\phi + u\sin\phi}{\sqrt{1-w^2}} \tag{5.25}$$

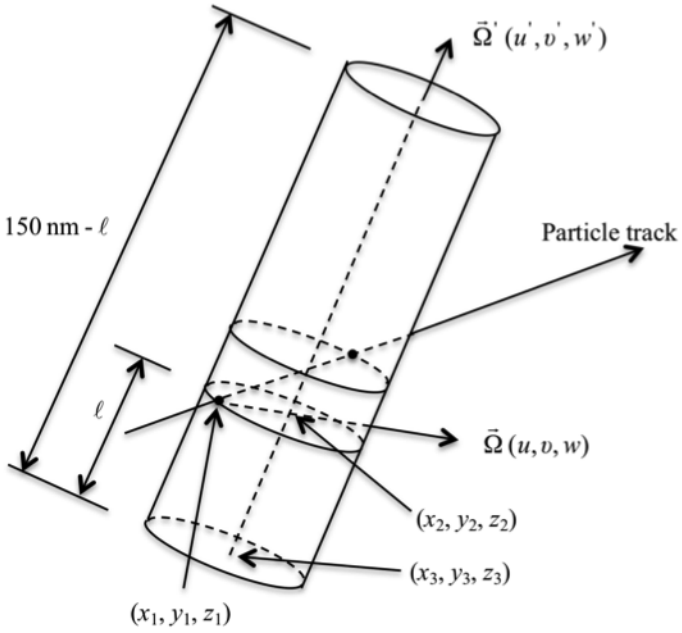


Figure 5.7: The geometric orientation of a CF rod with respect to the particle track intersecting it

$$w' = -\cos\phi\sqrt{1-w^2} \tag{5.26}$$

End point positioning of CF rods

To tally the energy deposited by the particle inside the CF rod, one must define the boundary of the CF. rod The boundary of the CF rod is defined by three parameters, the center of the base, (x_3, y_3, z_3), the direction vector of its axis, $\vec{\Omega}'(u', v', w')$, and the total length, which is 150 nm. Because the distances between (x_2, y_2, z_2) and (x_3, y_3, z_3) varies arbitrarily between 0 and 150 nm, x_3, y_3, and z_3, can be obtained as:

$$x_3 = x_2 - u'l \tag{5.27}$$

$$y_3 = y_2 - v'l \tag{5.28}$$

$$z_3 = z_2 - w'l \tag{5.29}$$

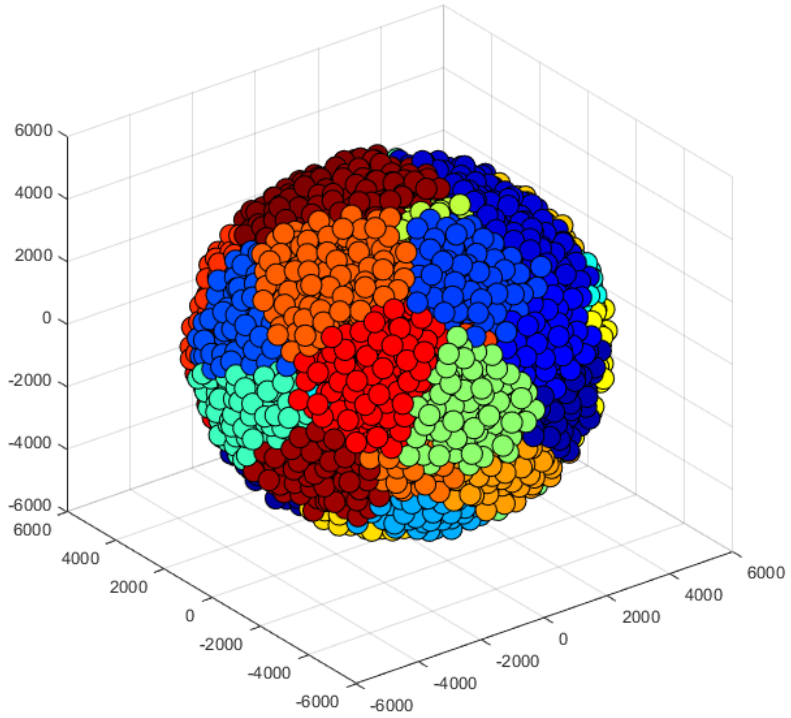


Figure 5.8: Chromosome Territories

where $l = (150nm) \cdot \mathcal{R}$ of which $\mathcal{R}$ is a random number between 0 and 1.

5.3 Chromosome territories

In addition to the other aforementioned structures, a recent advancement of our model includes the incorporation of chromosome territories for study as is seen in Figure 5.8. The code has been adapted to tabulate and identify interactions within CTs and between adjacent CTs.

The method by which CTs have been positioned is nearly identical to the method used for placement of CDs as described in section 5.1.1. A minimum distance threshold of 1 μm is specified for a sphere packing algorithm to ensure CTs are not allowed to be too small, and a total of 46 'spheres' are placed in the cell nucleus. The purpose is to position 46 points (corresponding to the centers of the spheres) within the cell nucleus to identify

the center of a CT. CDs are subsequently assigned to the CT it is closest to. Since the CT center points are assigned randomly, there is substantial variability in the number of CDs belonging to the different CTs.

This work is still in its preliminary stages. More investigation is needed to determine an appropriate assignment of chromosome numbers, CT locations, and DNA content within CTs. However, from a programming perspective, all of the coding has been set in place for differentiating damage based on CT information, including distinctions of chromosome aberrations. This effort will be reserved for future investigation as the model develops in the finer details of biological response.

CHAPTER 6

PRODUCTION OF DNA DOUBLE STRAND BREAKS

6.1 Methods

The next step is to take the energy deposition events recorded in various CF rods and convert them to the most relevant biological endpoint, the DNA double strand breaks (DSBs). Two major assumptions are made: First, the DNA molecule is assumed to be uniformly embedded in the CF rods. Second, the DSBs are produced solely by energy clusters. An energy cluster is defined as a cluster of energy deposition events that occur within a few nanometers from one another. With these two assumptions, one avoids the complexity of explicitly modeling the DNA structure. In other words, any energy cluster occurs inside a CF rod has a chance to produce a DSB. The probability for an energy cluster to produce a DSB depends on the size of the cluster, and it can be calibrated via the well known fact that one Gy of absorbed dose of any charged particle type produces approximately 40 DSBs per cell [1].

6.2 Intersection of Particle Tracks and DNA Structures

6.2.1 Intersection of energy depositions with CFrods

At this stage the cell nucleus model progresses to the identification of charged particle energy depositions to CF rods. Each CF rod has Cartesian coordinates corresponding to its endpoints locations, and also has an identifier denoting which CD number it belongs to or if it is within the IC. In essence the model is converting each CF rod into a unique detector region that has been randomly placed within the cell nucleus based on the positioning method discussed in Section 5.2.1. The CF rods are subsequently converted to cylindrical surface points through an open-source Matlab code [100], which are converted to a

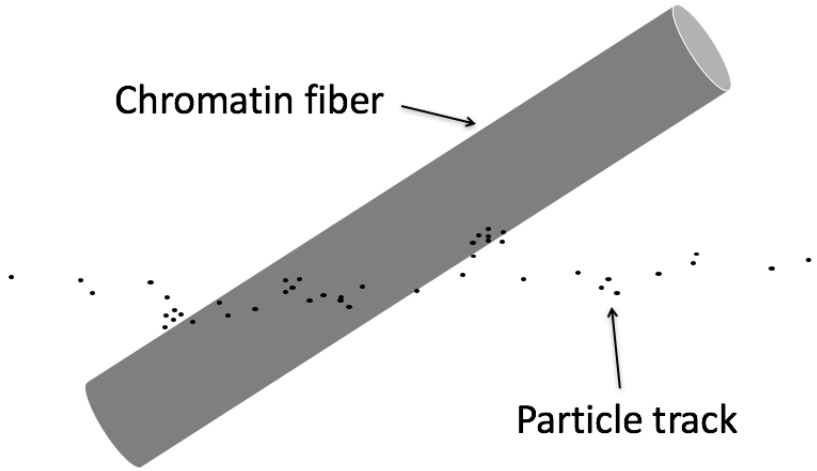


Figure 6.1: Intersection of particle track with CF rod

Delaunay-triangulated surface boundary. Another open-source Matlab code, inpolyhedron [101], is used for converting this surface boundary into a detector region to determine if charged particle events are located within the CF rod, which is demonstrated in Figure 6.1. This process is repeated for every CF rod, but only for energy deposition events that occur within the specific CD or IC region the CF rod belongs to.

Up to this point, the cell nucleus model as presented has been sufficiently computationally efficient. The production of charged particle track data using Geant4-DNA can be computationally expensive, but this work by itself is not novel. Several extra coding lines were added to the cell nucleus model to boost the efficiency of calculation, as the quantity of data being processed have introduced significant computing challenges, especially given the modest budget of the work presented. The knnsearch function in Matlab, previously used for separation of events within CDs versus within the IC, was further applied to each CF rod to filter out events that do not occur in the DNA-containing targets. This was done by performing the knnsearch to two points representing 25 percent of the length of the CF rod from the two ends of the CF rod. As the CF rod has a 150nm length, the two points are located 37.5 nm from the two ends of the CF rod. Any charged particle event occurring more than 40.4 nm from any of these two points was filtered out and excluded from the calculation. This step sped up the more intensive simulations by over an order of

magnitude.

6.2.2 DBSCAN algorithm for identifying energy clusters occurring inside CF rods

As mentioned in Section 6.1, to produce DNA DSBs one must first identify the energy clusters that occur inside the CF rods. The density-based clustering algorithm with noise (DBSCAN) was used for this purpose [102]. The input parameters used for DBSCAN were 2 nm as the radial distance, as it is the width of the DNA double-helix, and a minimum number of events of 3, essentially guaranteeing that at least 3 energy deposition events would have to be in very close proximity to be registered as an energy cluster. The rationale is that single base damage or SSBs would be incapable of causing DNA DSBs, and as such, the minimum requirement for production of DSBs is that at least a few energy deposition events would have to be in close proximity. Accordingly, sparse and isolated energy deposition events can effectively ignored. Figure 6.2 shows an artistic depiction of the effect of the the distance parameter in DBSCAN's clustering algorithm. Too small of a distance parameter effectively increases the noise of the simulation.

Output from DBSCAN allocates whether or not charged particle energy deposition events belonged to a cluster or not, and also provides a cluster identifier as to which cluster an event belongs to if multiple clusters are deposited within a single CF rod. The mean position of a cluster in terms of its centroid was identified, as were the total number of excitation and ionization events contained within the cluster, in addition to the total energy deposition within the cluster.

6.2.3 Method for recording energy clusters in a CF rod

Several counting methods were investigated regarding energy clusters within a CF rod, but in the end all energy deposition events were summed and all centroids were reported. For example, in the event that two clusters were formed within a single CF rod, the energy reported for the cluster would be summation of both clusters, with the code effectively

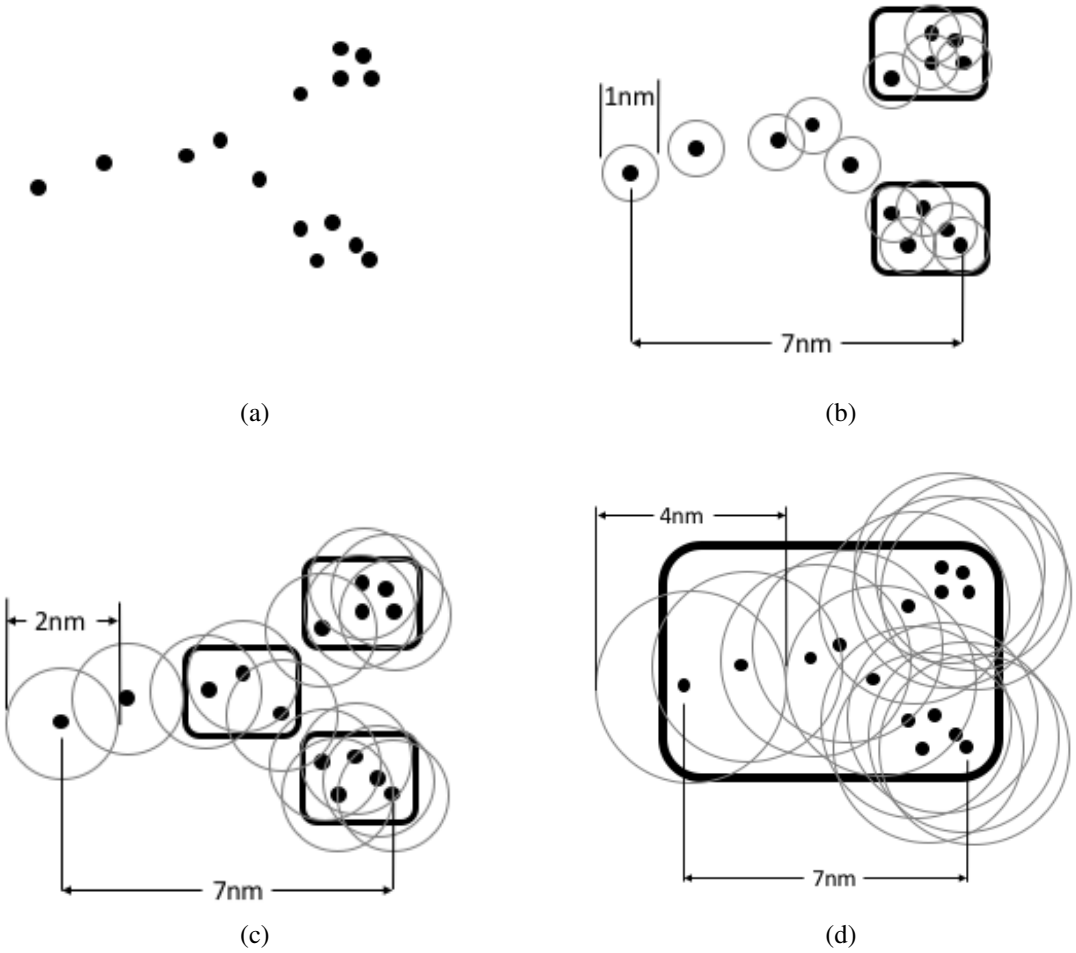


Figure 6.2: Illustration of the effect of the distance parameter in DBSCAN's clustering algorithm on a spatial distribution of energy deposition events: (a) the spatial distribution of the energy deposition events, (b) with the radial distance of 1 nm, (c) with the radial distance of 2nm, and (d) with the radial distance of 4nm.

reporting two clusters each containing the sum of their energies. Assessment of DSBs and misrejoining probability introduces a filter to ensure that clusters within a CF rod are not counted multiple times. The reason for allowing multiple cluster positions within a single CF rod is purposeful for later applications. Earlier investigations only looked at the largest cluster, but assessment of misrejoining probability dictated that a more inclusive method would be warranted, as it may be possible for a single CF rod to be multiply damaged and may possibly be involved in multiple misrejoining interactions.

6.2.4 DSB scoring method

Because only a small percentage of the CF is occupied by DNA, the probability for an energy cluster (occurring inside a CF rod) to produce a DNA DSB is likely to be small. In addition, it is important to differentiate the two types of DSBs, the simple double-strand break (sDSB) and the complex double-strand break (cDSB). This is because, as described in Chapter 2, the repair for sDSBs is rather straightforward (i.e. via NHEJ). The repair for cDSBs, on the other hand, are rather challenging. The probability for an energy cluster to produce a sDSB or cDSB is determined based on a statistical approach: It uses an energy cluster threshold of 75 eV as the step size with one or more random numbers depending on the total energy deposition of the cluster. The 75 eV threshold is based on previous studies indicating the possibility of DSB production at this energy level. Applying this energy threshold to electron irradiation and using 40 DSBs Gy^{-1} for calibration yield a 2.8% likelihood for a 75 eV cluster producing a sDSB. The Monte Carlo procedure is described as follows: The RNG is called once for a 75 eV cluster, twice for a 150 eV cluster, three times for a 225 eV cluster, etc. Since the chance for a 75 eV cluster not to produce a DSB is 97.2%, the chance for a 150 eV energy cluster to produce a DSB would be $1-(0.972)^2$ =5.52%, and that for a 225 eV energy cluster would be $1-(0.972)^3$ =8.17%, etc. The energy clusters wherein a single random number value meets the threshold criteria are counted as sDSBs, while clusters where multiple RNG values meet the threshold criteria are counted

as cDSBs. The above approach presumes that the spatial distribution of DNA inside a chromatin fiber is uniform. In this manner, larger energy clusters have a higher likelihood of producing DSBs while still accounting for uncertainties in the composition of chromatin structures.

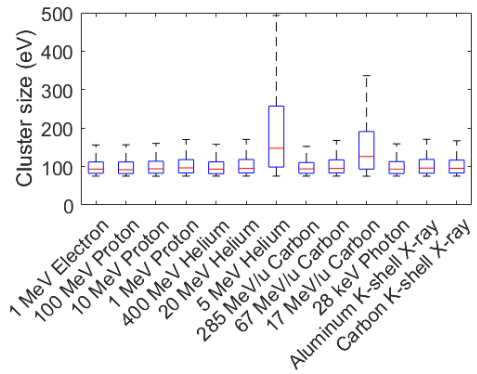
6.3 Model Validation and Results

6.3.1 To determine the optimal radial distance for DBSCAN for determining energy clusters

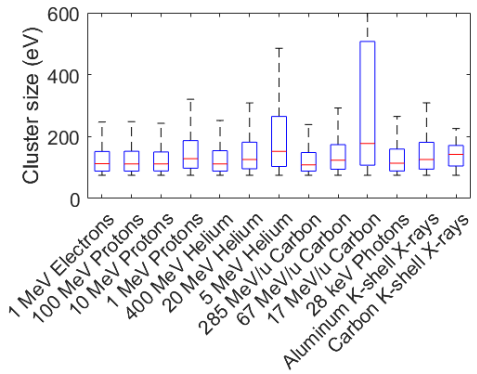
One of the challenges in this simulation study is to find the appropriate input parameters for DBSCAN for identifying energy clusters. DBSCAN operates with a requirement for two major parameters: radial distance and number of events. Both input parameters are straight-forward: the radial distance determines how close events must be to each other to be part of the same cluster, and the number of events determines the minimum number of events in proximity to be considered a cluster as is described in Figure 6.2. Some more complex parameters can be assessed, but for the purpose of our work this is sufficient: we assume that energy deposition events must be part of an energy cluster to be considered capable of producing a DSB. It is easy to see from Figure 6.2 how the clustering parameters influence the results from the model: too small of a radial distance makes the data appear 'noisy', while too large of a radial distance tends to give away the fine details associated with the energy clusters. The challenge is to find a metric that reasonably matches the observed results from cell irradiation experiments.

6.3.2 Detailed Analysis of Cluster Sizes and Pattern

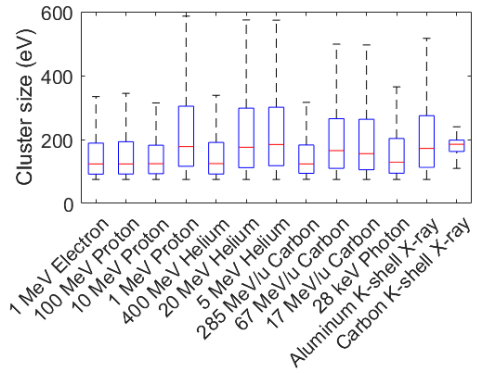
The method used to obtain the optimal radial distances for DBSCAN was based on the assumption that carbon K-shell X-rays would produce the most severe energy clusters, i.e. on the range of 280 eV. This method was performed by applying the DBSCAN clustering algorithm to a single track calculation output from Geant4-DNA and to output the total



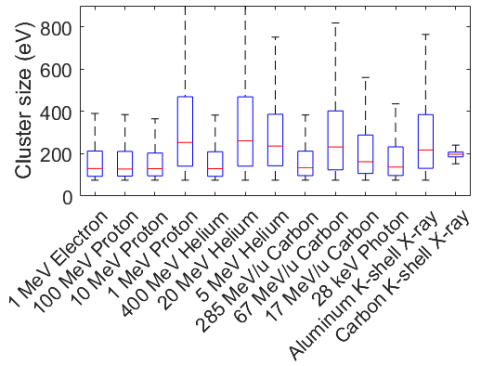
(a) 1nm radial distance



(b) 2nm radial distance



(c) 3nm radial distance

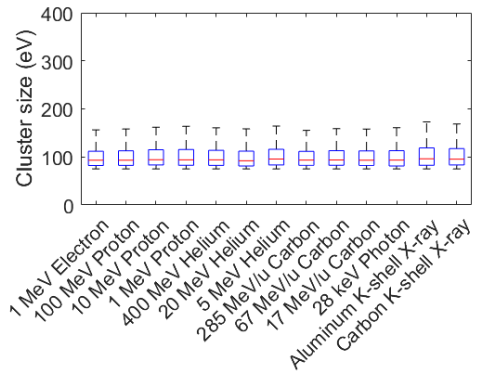


(d) 4nm radial distance

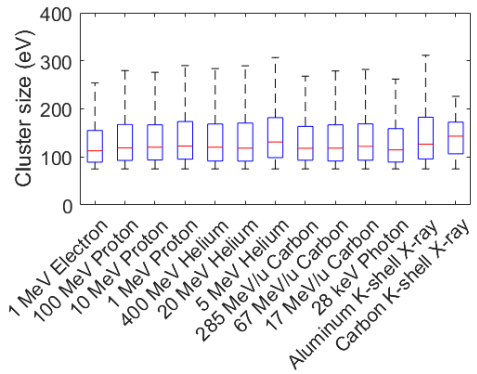
Figure 6.3: Comparison of cluster size distributions produced by DBSCAN for different radiation types and for the four radial distances, 1 nm, 2 nm, 3 nm, and 4 nm, respectively. Note: All clusters are greater than 75 eV

cluster information from the process. Figure 6.3 shows a comparison of cluster size (in eV) distributions produced by DBSCAN for different radiation types and for the four radial distances, 1 nm, 2 nm, 3 nm, and 4 nm, respectively. What can be clearly seen is that there is a substantial variation in cluster sizes among the different radiation types.

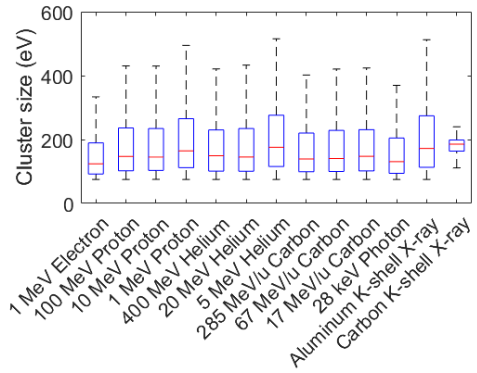
Among the four panels on display, Figure 6.3 shows that the radial distances of 2 and 3 nm give the greatest median cluster for the carbon K-shell X-rays relative to other radiation types. For the radial distance of 4 nm, the 1.5 keV aluminum K-shell X-rays produce larger clusters, likely due to its slightly more sparse deposition pattern resulted from the larger radial distance. On the other hand, Figure 6.3 (a) shows that the radial distance of 1 nm is obviously too tight to hold the entire 280 eV energy cluster (of the carbon K-shell X-ray)



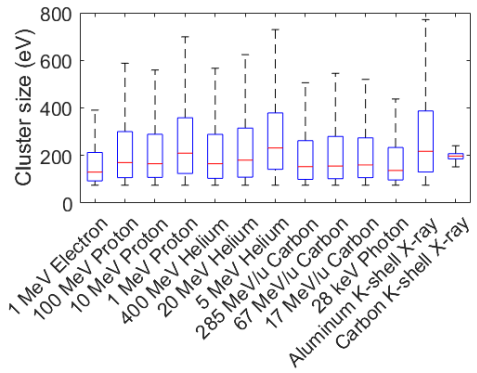
(a) 1nm radial distance



(b) 2nm radial distance



(c) 3nm radial distance

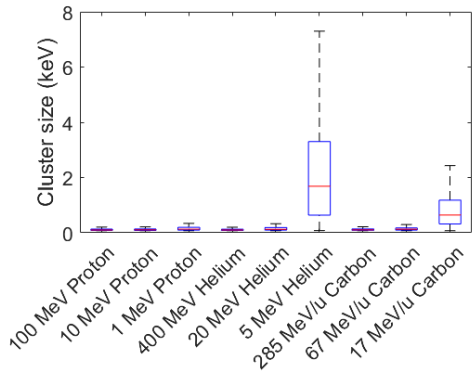


(d) 4nm radial distance

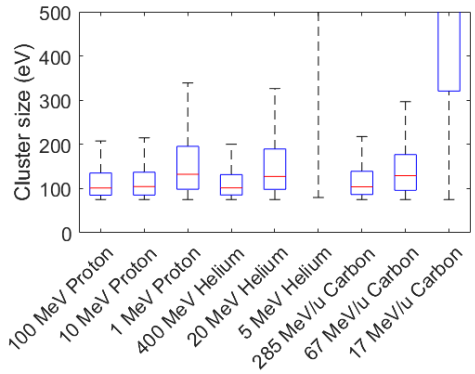
Figure 6.4: The energy cluster components that are contributed from the Knocked-out electrons for different radiation types and for the four radial distances, 1 nm, 2 nm, 3 nm, and 4 nm, respectively. Note: All clusters are greater than 75 eV

together consistently, making for the more sparse electron and photon energy depositions to be equally damaging relative to the carbon K-shell X-rays. Specifically, it is seen that, apart from 5 MeV helium ions and 200 MeV carbon ions, the median, 25th and 75th quartiles are basically identical. From the above analysis, it was concluded that the radial distances of 2 nm or 3 nm are about optimal to be used for the DBSCAN to identify and characterize energy clusters. In addition, these distances are consistent with the fact that the radius of DNA double-helix is approximately 2 nm. To simplify the study, we therefore decided to use the 2-nm DBSCAN results for the rest of the simulation study.

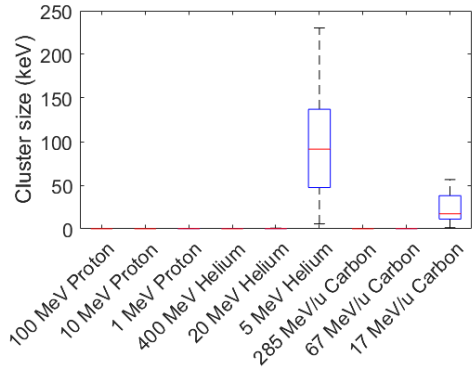
An additional analysis was taken by sorting out the two components of an energy cluster, i.e. the contribution from the knocked-out electrons and the contribution from the



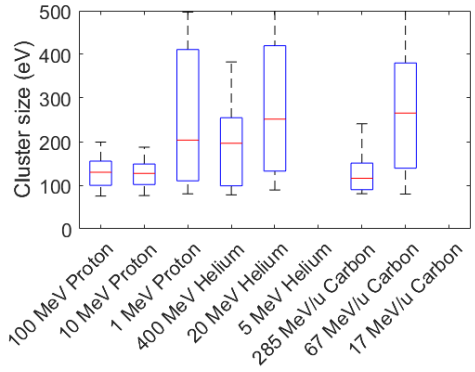
(a) 2nm radial distance



(b) 2nm radial distance, zoomed in



(c) 4nm radial distance



(d) 4nm radial distance, zoomed in

Figure 6.5: The energy cluster components that are contributed from the primary particles for different radiation types and for the four radial distances, 1 nm, 2 nm, 3 nm, and 4 nm, respectively. Note: All clusters are greater than 75 eV

primary particle. The cluster size distributions of these two components are respectively shown in Figures 6.4 and 6.5. By comparing with Figure 6.3, it is clear that the contribution from the knocked-out electrons is the major component of an energy cluster for most of the radiation types except 5-MeV helium ion. This can be explained by the fact that, among all the ions, 5-MeV helium is the only one that has a mean free path (in liquid water) approaching 2-3 nm, the radial distance used to identify an energy cluster. In other words, for low-energy ions (e.g. 5-MeV helium) with their mean free paths comparable to the radial distance that identifies a cluster, the soft-collision events (i.e. those that are $<$50 eV) are now mostly "connected" along the primary particle track and therefore are more likely to be identified as (or included in) an energy cluster.

Table 6.1: The percentage dose from the knocked-out electrons and the primary particles for the different radiation types. Note: the cluster was identified by DBSCAN using 2 nm as the radial distance

	By knocked-out electrons	% by primary particles	By clusters	By knocked-out electron clusters	By primary particle clusters
1-MeV Electrons			24%	24%	
Photons, 28 keV			30%	30%	
Photons, 1560 eV			54%	54%	
Photons, 286 eV			66%	66%	
Protons, 100 MeV	72%	28%	26%	35%	4%
Protons, 10 MeV	74%	26%	31%	38%	10%
Protons, 3 MeV	73%	27%	42%	47%	28%
Protons, 1 MeV	71%	29%	59%	59%	59%
Helium, 400 MeV	66%	34%	26%	37%	5%
Helium, 20 MeV	69%	31%	56%	54%	61%
Helium, 5 MeV	69%	31%	81%	73%	100%
Carbon, 290 MeV/u	70%	30%	33%	40%	18%
Carbon, 67 MeV/u	69%	31%	53%	50%	58%
Carbon, 17 MeV/u	71%	29%	70%	59%	99%

Table 6.1 shows the percentage of dose contribution that belongs to several categories based on the 2-nm radial distance of DBSCAN cluster analysis. Furthermore, the percentage of clustered dose varies substantially based on the radiation type. Interesting notes

can be seen regarding the clustered dose contributions from knocked-out electrons versus primary particle across radiation types, especially for the heavy ions with varying kinetic energies. It is readily observed that the vast majority of dose contributed from the primary particle is sparse in nature while the particle kinetic energy is high, and this value increases dramatically as the particle slows down.

6.3.3 Limiting factors for DSB yield

Our model restricts the priority of energy depositions to those occurring within CF rods and those that are densely clustered in a manner we believe to be relevant to DNA DSB production. The thought process behind this assessment, then, is heavily dependent on the proportion of a total dose to a cell nucleus that is clustered. Furthermore, our method also inherently defines the DSB yield as being proportional to the total clustered dose deposited in a cell nucleus: if all of the dose is clustered then all of it can contribute to DSB production, if half of the dose is clustered then only half of it can contribute to DSB production.

Additionally, on the per unit dose basis an increase in LET of a particle type would similarly lead to an increased likelihood of producing cDSBs but simultaneously reduce the total number of CF rods that would be intersected, reducing the DSB yield. The increase in cDSBs and their level of complexity by our model would increase with very high LET at a rate similar to the decrease in total DSB yield. Whether or not this is happening in the actual cell has not yet been explored, but it would add extra support for our statistical method of DSB tallying if correlations can be made.

Another way of describing this effect is to state that our model inherently limits the total DSB yield of a radiation type in a cell to a specific maximum value defined by the proportion of energy deposition that is clustered, as well as the total number of CF rods that are intersected by the particle tracks. Effectively, there is a theoretical maximum number of DSBs that could be produced in a cell nucleus by radiation, of any quality, for a given dose. This upper limit in our model is the total energy deposited in the cell for a given dose,

multiplied by the fraction of dose that is clustered, and further multiplied by the probability of an energy cluster to lead to a DSB and divided by the average energy of a cluster.

6.3.4 Distribution of energy clusters and DSB yields

The four columns of Table 6.2 show, respectively, the number of CF rods receiving energy depositions (i.e. "# of CF hits"), the number of CF rods receiving clustered energy depositions (i.e. "# of clustered CF hits"), the number of DSBs (i.e. sDSB+cDSB), and the number of cDSBs. The results were all normalized to per Gy per cell. As indicated, while there are large numbers of CF rods receiving hits, only a fraction of them receiving clustered hits that are relevant to the production of DSBs. These fractions vary greatly among the different radiation types, and are consistent with the values shown in Table 6.1. In general, the fraction is large for the high-LET particles (i.e. low-energy ions) and low for the low-LET particles (i.e. high-energy ions). The exception is noted, however, that the fraction drops for the two highest-LET particles, i.e. 5-MeV helium and 17-MeV/amu carbon ions. This exception is consistent with and can be explained by the "over-kill effect" of the RBE-vs-LET relationship of Figure 2.2.

As discussed in Section 6.2.4, the DSB yields were calculated and calibrated so that sparsely ionizing radiation would produce between 40 to 50 DSBs per Gy. This is reflected in Table 6.2 where the DSB yield for 1-MeV electrons is 42.2 DSBs per Gy. Because our model assumes that the DSBs can only be produced by the clustered hits, the variation of DSB yields among the different radiation types (i.e. column 3 of Table 6.2) follows a similar pattern as that of the clustered CF hits (i.e. column 2 of Table 6.2). In general, the DSB yields are higher for high-LET radiation particles. But the variation between high-LET radiation and low-LET radiation is at most a factor of 2 or so. This result is in good agreement with that observed experimentally.

Another interesting thing to note is that the ultrasoft X-rays of carbon and aluminum have higher yields of DSBs than other sparsely ionizing radiation types. This can partly

Table 6.2: Distribution of energy depositions in CF rods and DSB yields. All the numbers are normalized to 1 Gy of absorbed dose per cell

	# of CF hits	# of clustered CF hits	# of DSBs	# of cDSBs
1-MeV Electrons	3683	280	42.2	2.6
X-rays, 28 keV	2765	329	50.5	3.2
X-rays, 1560 eV	689	510	83.2	6.6
X-rays, 286 eV	1027	585	99.8	5.2
Protons, 100 MeV	3383	293	44.7	2.6
Protons, 10 MeV	1782	351	53.2	3.0
Protons, 3 MeV	877	447	66.4	4.1
Protons, 1 MeV	303	546	93.2	15.0
Helium, 400 MeV	2778	295	44.6	2.5
Helium, 20 MeV	576	555	85.6	7.6
Helium, 5 MeV	152	266	66.5	34.7
Carbon, 290 MeV/amu	1254	380	55.8	3.1
Carbon, 185 MeV/amu	1060	419	61.8	3.5
Carbon, 67 MeV/amu	844	533	82.1	6.6
Carbon, 17 MeV/amu	568	334	69.5	19.1

be explained by the proportion of energy depositions by these two radiation types in Table 6.1, wherein a very large proportion of their total energy deposition is clustered. While both ultrasoft X-rays are still considered sparsely ionizing, owing to their photon depositions, the higher clustering by the photoelectron means that a very large proportion of energy deposition is capable of clustering, as opposed to more sparse higher energy electrons and photons, wherein a large proportion of energy deposition is sparse and non-clustered. Table 6.1 shows this clearly in that the clustered proportions for 1-MeV electrons and 1560 eV X-rays are 24% and 54%, respectively, using 2 nm as the radial distance for cluster identification.

From the biological effect stand point of view, the most relevant result shown in Table 6.2 is the cDSB yield. This is because most of the sDSBs can be effectively repaired by NHEJ. Differences in how various particles and energies deposit their energy can be deduced from the ratio of DSBs that are simple vs complex. Electron cDSB to sDSB ratio can be used as a reference as sparsely ionizing radiation with other types of radiation. Protons at an energy of 100 MeV and helium ions at an energy of $400\,\mathrm{MeV}$ have nearly the same ratio of cDSBs to sDSBs as electrons.

As the heavy ion kinetic energy decreases, the proportion of cDSBs can be seen to increase, indicating an increase in the prevalence of complex damage as the particles' energy deposition becomes more condensed. 200 MeV carbon ions and 5 MeV helium ions have similar LET of approximately $100\,\mathrm{keV/\mu m}$, but the helium ions are seen to have a higher proportion of cDSBs than carbon ions at like LET. The same relationship can be seen between $800\,\mathrm{MeV}$ carbon ions, $20\,\mathrm{MeV}$ alpha particles, and $0.9\,\mathrm{MeV}$ protons, which shows how lighter ions have a tendency to have more densely clustered energy deposition at the same LET as heavier ions.

What is also significant about these values is that it shows quantifiable differences in damage to a cell versus DSB production. cDSBs have long been understood to be a more relevant factor than DSBs in terms of radiation-induced effect to a cell, but even so RBE

comparisons in various radiobiological models often allude simply to DSB production. It is significant that our model is able to differentiate between a sDSB and a cDSB.

It needs to be stated that this behavior is expected given how our model has defined the production of DSBs and cDSBs: we allocate the total energy within a cluster is proportional to the amount of energy in said cluster, and the production of a cDSB is tallied when multiple positive DSB counts are found in a single cluster. The higher LET particle types, as stated previously, are densely packed and frequently exceed $1\,\mathrm{keV}$ of total energy deposition within the cluster. We believe this approach to be reasonable for assessing DNA damage: nevertheless, we also set this parameter to be how our model would function.

6.3.5 Single track effects vs multiple track effects in causing DSBs

One of the applications of the cell nucleus model is for health physics with respect to low-dose radiation risk assessment where the biological effect is mostly caused by a single radiation track. The epidemiological data do not produce statistically meaningful results at dose levels lower than about 100 mSv [103]. A unique feature of our cell nucleus model is that the DSBS produced in the model are tagged with information regarding the origin of the subnuclear structure and location, and also retain information from Geant4 regarding the involved particles. The charged particle track outputs from Geant4 allows for tracking each energy deposition of primary particles and delta rays to their originating particle. The five particles depositing the most energy to each DSB is recorded: in this way, we can see how many independent particle tracks contribute to DSBs.

Table 6.3 shows the percentage of DSBs that are caused by multiple particle tracks. It indicates that DSBs are overwhelmingly the result of a single particle track. DSBs that have any minimal involvement of more than a single track are tallied without discriminating the proportion of energy contributed by originating primary particles.

Table 6.3: Percentage of DSBs caused by multiple particle tracks

	0.5 Gy	1 Gy	2 Gy	4 Gy
Electrons, 1 MeV	0.00%	0.01%	0.00%	0.02%
X-rays, 28 keV	0.00%	0.00%	0.03%	0.01%
X-rays, 1560 eV	0.00%	0.00%	0.00%	0.02%
X-rays, 286 eV	0.00%	0.01%	0.01%	0.01%
Protons, 100 MeV	0.01%	0.00%	0.01%	0.01%
Protons, 10 MeV	0.01%	0.01%	0.00%	0.02%
Protons, 3 MeV	0.00%	0.00%	0.01%	0.01%
Protons, 1 MeV	0.00%	0.00%	0.00%	0.00%
Helium, 400 MeV	0.00%	0.01%	0.00%	0.02%
Helium, 20 MeV	0.00%	0.00%	0.01%	0.01%
Helium, 5 MeV	0.00%	0.00%	0.01%	0.02%
Carbon, 290 MeV/amu	0.00%	0.00%	0.00%	0.01%
Carbon, 185 MeV/amu	0.01%	0.00%	0.01%	0.01%
Carbon, 67 MeV/amu	0.00%	0.00%	0.00%	0.01%
Carbon, 17 MeV/amu	0.00%	0.00%	0.01%	0.01%

6.3.6 Single track effects vs multiple track effects - misrejoining potential of different dose levels

After the model showed the prevalence of single radiation tracks being almost solely responsible for individual DSBs, some interest was generated on the proportion of potentially interacting DSBs that are a result of single radiation tracks. Essentially, what we wanted to deduce was how many potentially misrejoining DSBs are produced solely by a single primary particle and its associated delta rays, versus the potentially misrejoining DSBs involving multiple particle tracks. The interest for such an analysis is, in more practical terms, what possible α and β values of the LQ model could exist, and how our model can help link them to the single and multiple track effects.

Applying the DBSCAN algorithm to DSBs produced within the model with a radial distance of 2 nm, the output of clustered DSBs is further processed by applying DBSCAN one more time using 1 um as the radial distance to obtain the potential DSB misrejoining events. The results are reported in Table 6.4, which shows that as the dose is increased, the contribution from multiple particle tracks increases. This pattern is consistent for all particles and energies, which is expected as a higher density of DSBs will yield a higher likelihood for DSBs of adjacent particle tracks to interact.

High-LET charged particles are seen to have the bulk of DSB misrejoining events as single-track effect which is unsurprising given that the densely ionizing particle track can produce DSBs in close proximity, allowing for misrejoining to take place. This effect is shown in Table 6.4, i.e. the majority of DSBs in close proximity are due to single track effects, and the proportion of the effect increases at lower doses.

What seems to be surprising, however, is how many potentially misrejoining DSBs for sparsely ionizing radiation are due to single track effects. It is surprising because the low-LET radiation are thought to only produced isolated DSBs from the sparsely distributed delta rays of a few hundred eVs of energy. A possible mechanism responsible for this effect could be the high-energy delta rays (of a few keV) that are capable of producing

Table 6.4: Percentage of DSB misrejoining events caused by single radiation track with 1 um as the radial distance for DBSCAN analysis

	0.5 Gy	1 Gy	2 Gy	4 Gy
Electrons	52%	31%	18%	9%
Photons, 28 keV	41%	25%	16%	8%
Photons, 1560 eV	34%	23%	14%	8%
Photons, 286 eV	1%	1%	1%	1%
Protons, 100 MeV	36%	28%	15%	8%
Protons, 10 MeV	59%	39%	21%	11%
Protons, 3 MeV	67%	53%	35%	18%
Protons, 1 MeV	94%	80%	67%	49%
Helium, 400 MeV	48%	31%	16%	9%
Helium, 20 MeV	89%	71%	57%	37%
Helium, 5 MeV	97%	96%	89%	79%
Carbon, 290 MeV/u	60%	45%	29%	15%
Carbon, 185 MeV/u	68%	50%	36%	17%
Carbon, 67 MeV/u	83%	69%	51%	31%
Carbon, 17 MeV/u	94%	92%	81%	64%

multiple energy clusters in close proximity. The dose at which multiple-track effect begins to dominate is above 2 Gy, which is interesting to note as this exceeds the standard fractionations of 1.8 and 2 Gy per fraction most commonly prescribed in radiotherapy.

Assessment of heavy ions at higher energy and lower LET indicates they similarly show features of sparsely ionizing radiation types. Table 6.4 shows that at higher doses, the multiple-track effect increases for protons at 100 MeV and helium ions at 400 MeV, and that the effect is similar to that of 1-MeV electrons. The increase in the multiple-track effect is a result of the large number of particle tracks required for depositing higher doses with lower LET, thus increasing the likelihood for adjacent tracks to produce DSBs in closer proximity.

6.4 Impact of low-energy electrons on single-track effect

One of the most fascinating results we have noted in our model is the prevalence of single-track effects for possible DNA misrejoining by sparsely ionizing radiation types. This observation was entirely unexpected on our end, and much of the effort in addressing this was actually an attempt to minimize it as we originally believed it to be a coding error. An exhaustive analysis of our data confirmed that this behavior was not in error, although a better effort was made to explain it more logically.

Results from Table 6.4 show the impact of single-track effect on potential DSB misrejoining is a substantial fraction for all radiation types at doses as high as 1 Gy (except for 286 eV photons). This pattern is understood as fact for high-LET particles, but for low-LET particles it is somewhat unexpected. Several investigators have alluded to the importance of electrons 'at the end of their track' contributing to excess damage.

To investigate low-energy electron contributions, we looked at the common parent particle to potentially misrejoining DSBs caused by the single-track effect. In Section 4.3.1 we detailed the selected outputs from Geant4-DNA for each energy deposition of the particle track, which includes the primary particle (the original), the identifier of the current

particle, and the direct parent particle (i.e., for a delta ray, which particle caused it). Hence, the data includes a particle's family tree, so in the case of single-track effects the nearest common ancestor can be extracted from the data. Table 6.5 shows the results of this assessment, indicating the percentage of nearest common particles with energies below 1 keV, between 1 keV and 100 keV, and above 100 keV.

Table 6.5: Nearest common parent particle kinetic energy for single track effects

	<1 keV	1 keV - 100 keV	>100 keV
1-MeV Electrons	59%	21%	20%
X-rays, 28 keV	65%	35%	0%
Protons, 100 MeV	56%	12%	32%
Protons, 10 MeV	50%	11%	39%
Protons, 3 MeV	40%	5%	55%
Protons, 1 MeV	15%	1%	84%
Helium, 400 MeV	53%	13%	34%
Helium, 20 MeV	26%	6%	68%
Helium, 5 MeV	18%	1%	81%
Carbon, 290 MeV/amu	41%	11%	48%
Carbon, 185 MeV/amu	36%	10%	54%
Carbon, 67 MeV/amu	27%	7%	66%
Carbon, 17 MeV/amu	27%	8%	65%

The last data column in Table 6.5 for parent particles above 100 keV is expected, as higher LET particles will have a more spatially connected track. The first data column in Table 6.5 shows that there are many potential DSB misrejoining events entirely caused by electrons below 1 keV. For the single-track effect, delta rays below 1 keV are causing multiple DSBs in the cell nucleus at an unexpectedly high rate across all radiation types. This pattern is the most pronounced in sparse radiation types. With the exception of 28 keV photons, all other radiation types have a small contribution in the intermediate energies between 1 keV and 100 keV.

CHAPTER 7

FORMATION OF CHROMOSOME ABERRATIONS AND CELL DEATH

Assessment of cellular endpoints through application of our model is one of the more challenging elements from a validation perspective. While the efforts of the prior chapters on calculating the spatially distributed DSBs are computationally exhaustive and time-consuming, translating those results into biologically observable entities (e.g. DSBs, chromosome aberraitions, and cell survival curves) involves a substantial amount of speculative work. Validation of energy loss, DNA density within various subregions of the cell nucleus, or even consistency of cluster parameters is logical and has a veritable certainty, but the translation of physical quantities into biological endpoints (involving various repair pathways) introduces large uncertainties.

The main strategy behind the simulation study is to combine the detailed modeling of subregions of a cell nucleus with the radiation track structure data to allow quantitative analysis of various types of radiation-induced biological endpoints, i.e. DSBs, chromosome aberrations, and cell survival curves. In other words, the ultimate goal of the study is to be able to properly convert the physical quantities of radiation, i.e. radiation dose and energy cluster, to the biological endpoints, i.e. DSB, chomosome aberration, and cell survival curve so that the variation of spatially distributed energy clusters can be used to successfully explain the RBE of the various radiation types.

7.1 Methods

Several methods are detailed in this chapter that show how we translate spatially distributed DSBs into a statistical analysis method for assessing cell survival. This statistical analysis method is largely absent in the literature for other radiobiological models.

We first look at the different possibilities for repairing and misrejoining of DSBs. Ow-

ing to distinctions in the subregions of the cell nucleus, we observe distinctions in the likely forms of misrejoining that can take place between DNA DSBs based on what CDs the CF rods belong to, how close DSBs are to each other, and how many DSBs are clustered together. From this point we can look at how prevalent different forms of chromosome aberration (as a result of DSB misrejoining) are for different radiation types, and we can begin to perform a statistical assessment of how frequently these aberrations occur.

Following a proper assessment of the statistical parameters of a given radiation type, we subsequently are capable of simplifying our cell nucleus simulations into a statistically-correlated Monte Carlo code that can be performed very rapidly. This was essential for performing our cell survival calculations.

7.1.1 DBSCAN analysis for calculating potential DSB misrejoining probability

The misrejoining probabilities for DSBs has been performed by assessing the complexity of DSBs, proximity of interacting DSBs, and the origin of interacting DSBs. This stage is performed by applying the DBSCAN clustering algorithm on the DSBs as defined in section 6.2.2. The minimum cluster size is defined such that any DSBs in close proximity are identified for possible sources of misrejoining. The DBSCAN radial distance is varied so as to determine the model's optimal settings to match experimentally observed parameters. Literatures published previously indicate that 1 um is an appropriate value to be used for this purpose [papers by Chen and Sachs].

The outputs from DBSCAN identify the quantity and origin of DSBs in proximity to each other. In this way, our model aims to differentiate the types of misrejoining and aberrations that are produced by ionizing radiation of varying types and quantities. This is possible as the origin of DSBs is preserved, and this informs on whether the CF rods are within the same CD or in the IC, if the DSBs are simple or complex, and how many such CF rods are involved. In addition, the originating charged particle event identifier is tagged to the DSBs. Each DSB itself has stored with it the original event identifier from Geant4,

and thus it is possible to identify if DSBs are caused by a single charged particle event or multiple, and also it allows for identifying interactions of DSBs that are caused by a single particle tracks or an interaction of multiple particle tracks.

7.1.2 Distinction of chromosome aberration types

It is important to note that the specific subregions (i.e. CD, IC and CF rod) of DSBs in proximity allows for the assessment of complexity of DNA damage. For example, misrejoining of DSBs occurring in the IC region of two different CTs is likely to cause interchromosomal aberrations (e.g. dicentric, acentric, and translocations); whereas misrejoining of DSBs occurring in the same CT is likely to cause an intrachromosomal aberrations (e.g. centric ring and acentric fragment). In addition, misrejoining of two DSBs occurring in the same CD is likely to result in the deletion of a small DNA segment ($<$ 1 Mbps), which in turn, will be recognized as a micronucleus formation during mitosis. While misrejoining of DSBs are likely the major cause of chromosome aberrations, one must not forget that an unrejoined DSB can also result in the deletion of a large DNA segment, i.e. a deletion-type chromosome aberration. In other words, by analyzing both the DSB types and locations one can deduce different types of DNA damage at the chromosome level.

The actual probability of misrejoining of nearby DSBs likely varies based on the interplay of two major factors, the distance between the DSBs (i.e. the proximity factor) and the type of the DSBs (i.e. the repair kinetic factor). The importance of the proximity factor is obvious in that the misrejoining probability must increase as the distance between the DSBs decreases. The type of DSBs is also important because repairing a DSB is highly dependent on the complexity of the DSB. For example, an isolated sDSB is usually rejoined via NHEJ pathway in approximately 30 minutes. As such, two nearby sDSBs only have a relatively short time window (and therefore small probability) for misrejoining to occur. On the other hand, cDSBs tend to linger around for a long time, and therefore, have higher probability for misrejoining to occur.

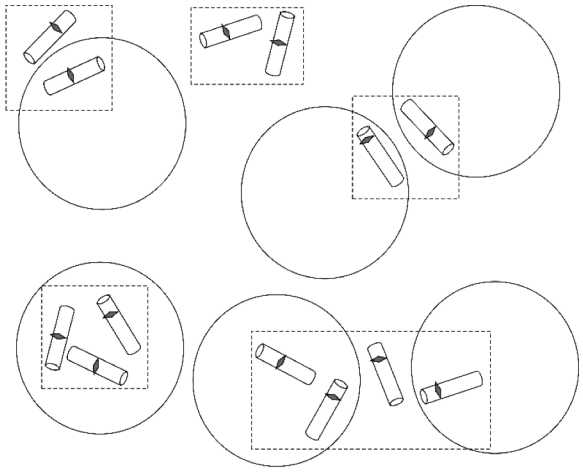


Figure 7.1: Possible scenarios of misrejoining of nearby DSBs

The starting point for our simulation of cellular endpoints is once again in the application of the clustering code DBSCAN for the determination of how close different DSBs are to each other, as well as how many are in sufficient proximity so as to possibly misrejoin. It is generally understood that loose DSB ends can migrate in the cell nucleus for a distance of as much as 1 um so as to misrejoin with another DSB.

Figure 7.1 shows categorically the various scenarios of misrejoining of nearby DSBs. The specific CF rod associated with each DSB is noted regarding which subregion it belongs to, be that two distinct CDs, the IC, or some combination that may possibly include more than 2 CF rods. We also note that the probability of certain scenarios of misrejoining by the NHEJ pathway may be entirely excluded from different combinations, while others may be more prevalent. Additional information is recorded regarding whether CDs belong to the same chromosome.

We have identified a few combinations of potential DSB misrejoining within our model. Part of our reasoning in distinguishing CF rods within CDs versus the IC is to acknowledge a possibly higher misrejoining probability for CF rods within the IC, as they are more exposed and more mobile within the cell nucleus: there is a higher likelihood they can misrejoin than for a CF rod contained within a more densely packed CD.

Multiple DSBs within a CD are similarly quantified separately. Misrejoining from

within a CD should only result in a loss of a relatively short fragment of less than 1 Mbps of DNA. Hence, we believe it is likely that small micronuclei observed in fluorescence microscopy is due to intra-CD misrejoining of DSBs.

DSBs of adjacent CDs are also noted to have a possible misrejoining event. These can introduce some level of complexity: if the CDs are within the same CT, they can lead to intrachromosomal cross-arm translocations, acentric rings, or larger micronuclei. If two CDs are from different CTs, they can lead to interchromosomal translocations and dicentric rings.

The type of chromosome aberration can get really complicated when more than two DSBs are in close proximity: the combinations of possible misrejoining get considerably more complex. In very high-LET radiation types as well as for very high dose levels, these types become more prevalent. Identifying the likelihood and probability for these types of chromosome aberrations and cell death becomes much more complex in such cases.

7.1.3 Probability of DSB misrejoining

At an early stage in the study, it is required to apply some simple metric to check if the model is at least headed in the right direction. It would not be feasible to have any level of confidence in the model without an early test, and so we took the approach to broadly apply an assumption that any type of DSB misrejoining would lead to a chromosome aberration, which in turn, has a chance to kill the cell. While trial and error factored into how we eventually settled on our calculation method, much of it was a more cognizant thought process regarding how DSB misrejoining would eventually pan out. From our perspective, the clustering algorithm applied to DNA DSBs made logical sense for producing chromosome aberrations: DSBs close to each other would have some chance of misrejoining. Therefore, an increase in the number of DSB clusters should proportionally increase the chance of producing chromosome aberrations.

Chromosome aberrations were identified in our model by applying the assumption that

DNA DSBs can have some free range to migrate before being identified and restituted by cellular repair mechanisms. In the literature, this value has frequently been cited as $1\,\mu\mathrm{m}$. We understood that it is likely for most isolated sDSBs to be properly repaired without consequence, and this was the reason for applying a random probability for each chromosome aberration to be lethal. Furthermore, we also identified there to be a much higher likelihood for a very large cluster of DSBs in proximity to lead to lethal chromosomal aberration than only two DSBs in proximity, and care was taken into how to factor in a higher likelihood of cell death from such an occurrence.

We applied a uniform cell killing probability to our cell nucleus model for DSBs in proximity: any events close by, irrespective of type, would have some small chance of killing the cell. To take such an approach still requires some assumptions about cell killing. As discussed previously, the level of complexity for misrejoining increases when more than two DSBs are found nearby; this occurs very frequently in the clustering analysis of DSBs for the model across all radiation types as dose escalates. In the case of two involved DSBs, we have chosen to assess there as being a single possible misrejoining event involved.

In the case of three or more involved DSBs, we have chosen to assess there being a single possible misrejoining event for each involved DSB: for three DSBs we count three possible misrejoining events, for four we count four possible misrejoining events, and so on. There are a lot more restitution options for three DSB ends, and the number increases dramatically beyond that. The simple math behind why the multiple combinations becomes chaotic is fairly simple: an increasing dose will increase number of DSBs, which given the limits of the volume of the cell nucleus, will increase the proximity of DSBs dramatically enough. Increasing the DSB complexity by a factorial combinatoric approach will bias our cell survival calculation beyond even the linear-quadratic model well beyond any sensible, experimentally observed pattern. We do acknowledge that the choice to limit 2 DSBs in proximity to a single possible misrejoining event may not be the best option, but as further results will highlight, this approach is suitable for our model at this point.

Challenges of assessing DSB misrejoining probability and complexity of chromosome aberrations

It was understood that even the simple choice of how to handle misrejoining probability for two DSBs in proximity is challenging enough as there are 4 loose ends involved with two DSBs. In practice, for four loose DSB ends the restitution options are correct repair, incorrect repair with one end, and incorrect repair with another end. Three possible choices for misrejoining exist, and this with the assumption that all ends are rejoined in some fashion. In the case where the two DSBs are from the same chromosome but in different arms, that would mean one result is no effect, one result is an acentric ring, another is a translocation, and another can lead to two lost fragments. In a case where the ends are from two different chromosomes, the combinations are no effect, translocation, multicentromeric lethal aberration, and two lost fragments.

Furthermore, another question is at what proportion would such rejoining events occur, and if there is any dependence on other factors, such as the complexity of a DSB, for certain events to be more prevalent. In addition, in situations where three DSBs are found to be in close proximity, we have decided to count it more complex, as well. Some bias is hence factored into our presented cell survival calculations owing to this determination of misrejoining probability, and this bias may not be what is naturally occurring in the cell. Assessments of cell survival in comparison to experimental data should help to clarify how pertinent these assumptions are, and adjustments can be made to correct any found discrepancies.

While the flowchart of our simulation code is structured to accommodate such details, for now such specifics are simply beyond the scope of this study. In other words we have decided to forego the more detailed elements of the analysis to target some tangible quantities that are still of benefit for correlations. Further advances to the model that may involve the interconnectedness between CDs within a chromosome territory can certainly be made, but an algorithm of a different nature would be needed than has already been coded into

our model. The strength of our model is that it has the ability to statistically project many of these areas, but we leave that to future work.

7.1.4 Monte Carlo method for simulating chromosome aberration and cell death

Regarding work that has taken place up to this point, it should be noted that we have taken a combination of the data from the charged particle tracks, the subregions within the cell nucleus, and the intersections of the two. We have also recorded substantial data regarding the possible DSB misrejoining events that can take place. However, work up to this point is computationally expensive and for practical purposes, very time consuming. A more efficient simulation of the cell nucleus is possible with the data collected through methods of the prior chapters.

Probability distribution functions of DSB misrejoining events

The data collected up to this point (on the per cell and per unit dose basis) include: the total number of DSBs, the type of each DSB (i.e. sDSB or cDSB), the spatial coordinates of the DSBs, and the subregion that each DSB belongs to. Our premise is that DSB misrejoining events form chromosome aberrations, which in turn, is mostly responsible for cell death. To assess the cell death (or survival) probability, one first converts the above data into the probability distribution function (PDF) of the number of potential DSB misrejoining events for each cell. As an example, Figure 7.2 shows such a PDF for 4 Gy of 1-MeV protons.

Cumulative distribution functions of DSB misrejoining events

By integrating the PDF of potential DSB misrejoining events shown in Figure 7.2, one obtains the cumulative distribution function (CDF) as shown in Figure 7.3. The importance of the CDF in Figure 7.3 is that it can be used directly with a uniform random number generator to determine the number of potential DSB misrejoining events in a simulated cell: the y-axis value as shown corresponds to the uniform random number, and the x-

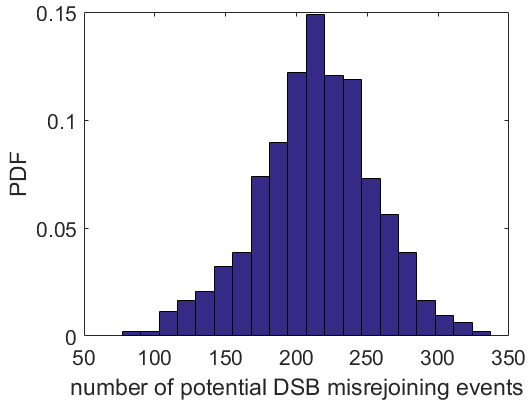


Figure 7.2: The probability density function (PDF) of the number of potential DSB misrejoining events per cell for 4 Gy of 1-MeV protons

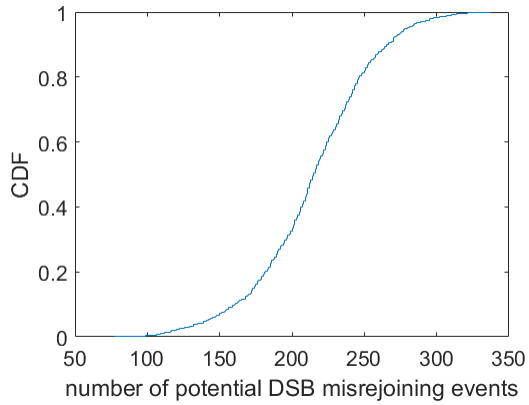


Figure 7.3: The cumulative distribution function (CDF) of the number of potential DSB misrejoining events per cell for 4 Gy of 1-MeV protons

axis value that matches would be the number of potential misrejoining events within the statistically-based cell nucleus.

The expectation value for the number of potential DSB misrejoining events in a cell nucleus can thus be predicted based on the spatially-based simulation data of DSBs.

Best fit assumptions for distributions

For higher doses of radiation and more densely ionizing radiation types, it is reasonable to expect that the normal or Poisson distribution may match well with the PDF of the number of potential DSB misrejoining events per cell: the quantity of DSBs per cell and

their proximity will allow for a many DSB misrejoining events to occur in each cell. For low-LET radiation types and low doses, where many of the cells will not register even one DSB misrejoining event, a normal distribution tends not to work well: an empirical assessment of potential DSB misrejoining events makes for a more consistent data set. This is performed through application of an empirical cumulative distribution function from the cell irradiation history taken from the model, with the prevalence being matched identically to the cell survival simulation. It is important to note that this works sufficiently given that the likelihood for a single DSB misrejoining event killing a cell is still very low in our simulations.

Expansion of the PDFs and CDFs beyond DSB misrejoining events

The rationale for use of a Monte Carlo approach for obtaining the number of potential DSB can further be expanded to other parameters for categorizing the types of likely misrejoining events. For example, we can take a look at all the DSB clusters within the cells and note the typical number of involved DSBs per cluster, taking note of the typical spread of such a cluster.

Given the number of potential DSB misrejoining events per cell that are determined from applications of Figure 7.3, for a given radiation scheme we can further break down the number of misrejoining interactions into the expected clusters of DSBs. From clusters of a said size, say three DSBs, we can also perform a statistical breakdown of the proportion of involved CDs and how many CF rods would likely be from the IC, as well.

Hence, from our spatially correlated cell nucleus model, we can extract statistical parameters to determine method for simplifying our cell simulations into a statistical pattern. This approach greatly simplifies more exhaustive simulation, and yet still keeps the original framework of the Monte Carlo-based simulation model. That is, this approach is nonetheless a cell-by-cell simulation and not a generalized analytical calculation. In effect, this is most certainly a Monte Carlo cell nucleus model, with the approach essentially being a full-

scale reduction of the input framework of the charged particle tracks and the subregions of the cell nucleus.

7.1.5 Cell survival curves

Cell survival calculations were made assuming that any form of DSB misrejoining could lead to chromosome aberration and cell death. While basic and elementary, especially given the extent to which different subregions of cell nucleus have been modeled into the code, the results from this approach yielded good agreement with expected patterns for cell survival. We adjusted several parameters to test different cell killing likelihoods based on different types of DSB misrejoining (or chromosome aberration), but the simple approach produced good agreement with the general behavior for cell survival calculations as was expected for this stage of work.

For simulation of cell survival, the method is straight-forward. For each potential DSB misrejoining event in a cell nucleus identified in section 7.1.4, a single RNG value is generated. If the RNG value meets a threshold criteria, the DSB misrejoining event is identified as lethal and the cell is killed. A cell with 20 potential DSB misrejoining events would subsequently have 20 independent RNG values generated, and any RNG value above a certain threshold would be counted as a cell killing event. Hence, if the highest value among the RNG values was 0.94 and a threshold was set as 0.95, the cell would survive: if the threshold was set at 0.90, the cell would have been killed. This process is repeated for as many simulated cells as needed until the result is converged to within a preset statistical uncertainty.

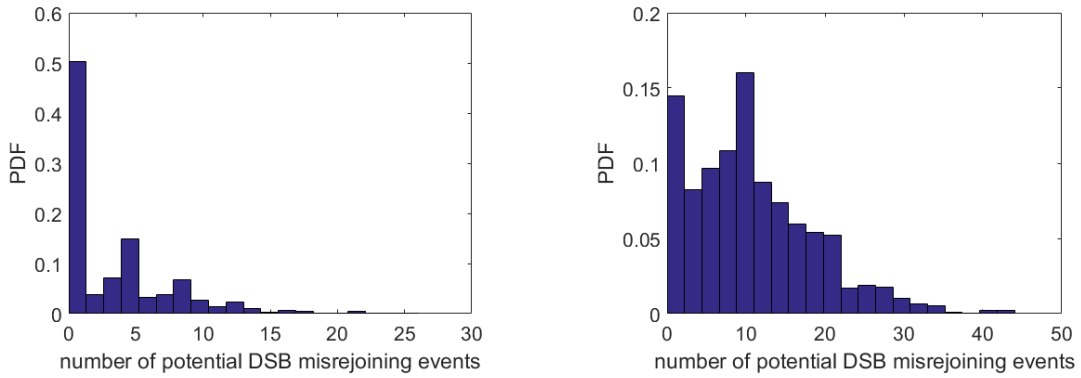
While this is how the cell survival curves are produced in this study, it should be noted that future work aims to improve on results using more complex parameters as highlighted in section 7.1.4. One of the purposes for distinguishing between different DSB misrejoining types in our model was so that some biological adjustment parameters could be performed within the model for specific cell lines and conditions.

Up to this point we have discussed that the primary variable in our simulations has been the radiation type and dose quantity: in the long-run, we intend for the dose rate and cell type to be additional variables, as well. To implement either one of these two variables, one would need to add the relevant biological response to the model. For example, for the dose rate effect one can add the kinetic factors for DSB repair to the simulation. One may use two different exponential decay functions (one for fast repair and one slow repair) to simulate the repair kinetics for sDSB and cDSB, respectively. As such, under low-dose-rate condition many sDSB would have been repaired during the course of irradiation. Accordingly the number of sDSBs would drop significantly compared to the high-dose-rate condition. The smaller overall number of DSBs averaged over the total irradiation time would, in turn, result in a smaller number of DSB misrejoining events and lower chance of cell death. For the cell type effect, one will need to adjust the model so that the biological response would have a strong correlation to the different cell types, and by adjusting the different CDFs and thresholds, our model would respond differently to radiation. In effect, the response of the model would be adjusted with a lower cell killing threshold in order to behave like a cell line that is more radiosensitive. The actual structural makeup of the cell nucleus in the cell simulation could similarly be adjusted to include more or less DNA with different CF rod densities, as well. In any case, the presented data represents only one such cell line.

7.2 Results

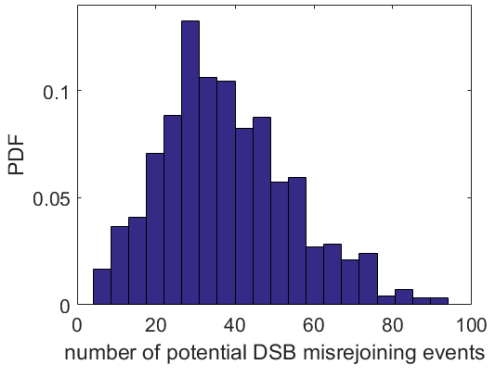
7.2.1 PDF of the number of potential DSB misrejoining events

According to the method discussed in Section 7.1.4, variations were noted for the number of potential DSB misrejoining events tallied in the cell irradiation simulations for the different radiation types and doses. The frequency of higher quantities of potential DSB misrejoining events per cell increased with increasing dose for all different types of ionizing radiation, which is expected. In addition, the number of potential DSB misrejoining events

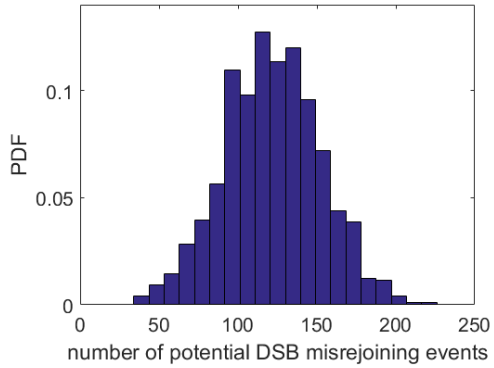


(a) 0.5 Gy

(b) 1 Gy



(c) 2 Gy

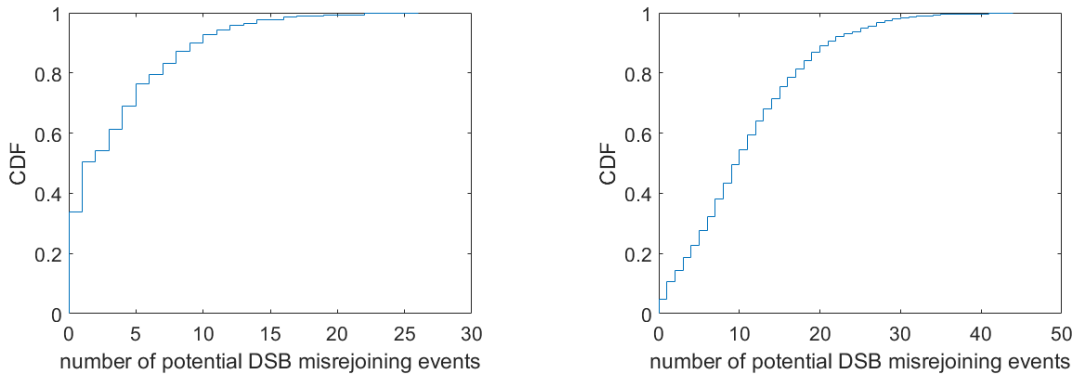


(d) 4 Gy

Figure 7.4: PDFs for the number of potential DSB misrejoining events per cell for 17-MeV/u carbon ion for the four different dose levels: 0.5 Gy, 1 Gy, 2 Gy, and 4 Gy

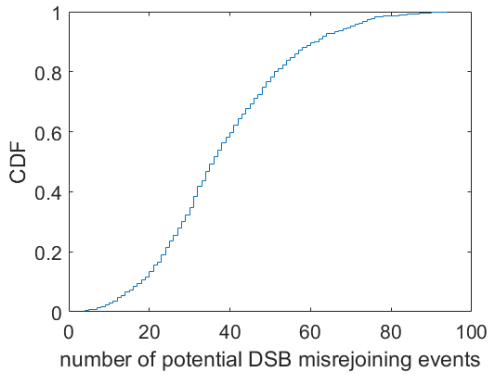
per cell tends to increase for higher LET radiation types, which is also expected. Figure 7.4 respectively shows the PDFs for the number of potential DSB misrejoining events per cell for 17 MeV/u carbon ions for each of the four dose levels.

There is a large variation in the number of potential DSB misrejoining events per cell within each of the four dose levels shown in Figure 7.4, indicating a large variation in damage received per cell even at the same dose level. For the 0.5 Gy dose quantity, a large proportion of cells are noted to have zero or very few misrejoining events per cell. At the 1 Gy dose range, the PDF somewhat resembles a Poisson distribution with a non-negligible portion encompassing zero misrejoining events per cell. A noted overlap of misrejoining events per cell is present across different dose ranges, indicating the damage to a cell by

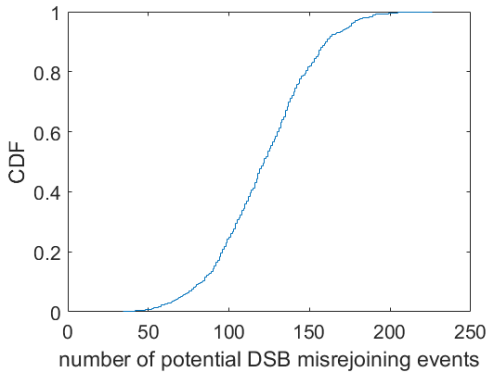


(a) 0.5 Gy

(b) 1 Gy



(c) 2 Gy



(d) 4 Gy

Figure 7.5: CDFs for the number of potential DSB misrejoining events per cell for 17-MeV/u carbon ion for the four different dose levels: 0.5 Gy, 1 Gy, 2 Gy, and 4 Gy

ionizing radiation is difficult to account for even under controlled circumstances.

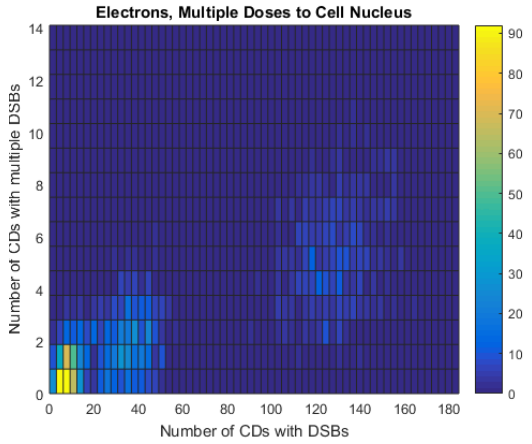
Figure 7.5 shows the empirical CDFs for 17 MeV/u carbon ions, which is used in the calculations for the number of misrejoining events per cell and shows more clearly the range and likelihood for misrejoining events per cell. One of the benefits of using the CDF instead of a curve fit is that the yield of misrejoining events is directly inferred from the simulations. Different statistical fits applied to the PDF of figure 7.4 tend to have issues where the probability of zero misrejoining interactions per cell is non-negligible.

7.2.2 Intra-CD micronucleus formation

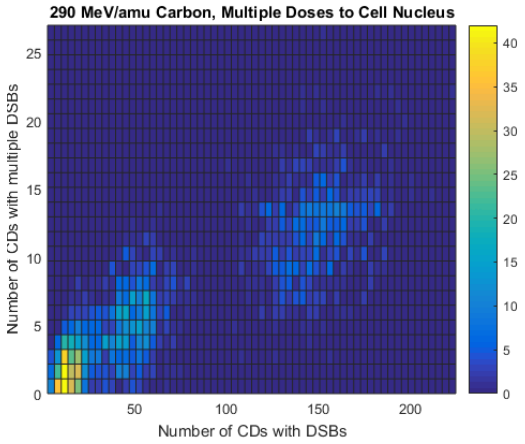
A specific endpoint that our model is uniquely able to produce is the simulation of micronucleus production. One type of DSB misrejoining event that we early identified as being unlikely to directly kill a cell following irradiation is that between CF rods within the same CD. Within a single CD, misrejoining events would most likely lead to a deletion of up to the entire 1 Mbp of the chromatin in the CD, but would be much less likely, owing to the nature of chromatin organization, to lead to cell death. Exclusion of such a relatively small fragment of DNA has been linked to the release of a small micronucleus following mitosis. Other sources for micronucleus formation are still possible, but within the context of our simulations, this type of intra-CD misrejoining is likely to produce DNA fragments with sizes ranging from a few kbps to 1 Mbps that can be quantitatively identified using pulsed-field gel electrophoresis (PFGE) technique. A very simple count with the model is to tally the number of CDs that feature multiple DSBs to different CF rods. Direct comparison of the relative ratio of such intra-CD DSBs can be used to compare the expected micronucleus yields of different radiation schemes.

The results shown in Figure 7.6 display the number of CDs containing DSBs on the x-axis and the number of CDs containing multiple CF rods with DSBs on the y-axis. Figure 7.6 intends to show the likelihood of different particles at specific kinetic energies of creating intra-CD micronuclei based on how likely a radiation type is to cause multiple DSBs within CDs. In all of the plots, data were obtained based on four dose levels: 0.5 Gy, 1 Gy, 2 Gy, and 4 Gy per cell. What is apparent is that more spatially-connected, high-LET forms of ionizing radiation have an increased likelihood of producing intra-CD micronuclei. The variance is lower for the densely ionizing radiation types, and the relative yields are much higher.

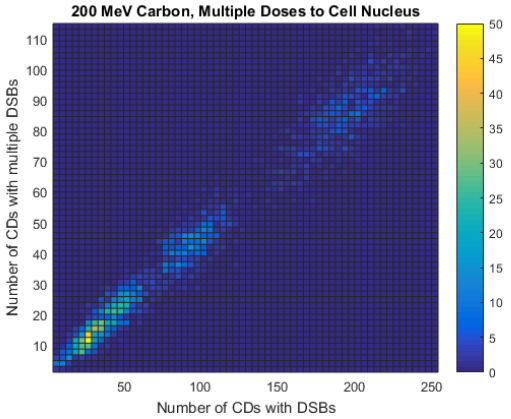
For the low-LET radiation types (i.e. 1-MeV electrons and 290 MeV/u carbon ions), Figure 7.6 shows that separate regions can be seen towards the high x-axis values corresponding to the doses of 2 Gy and 4 Gy, whereas there is no such separation for the low



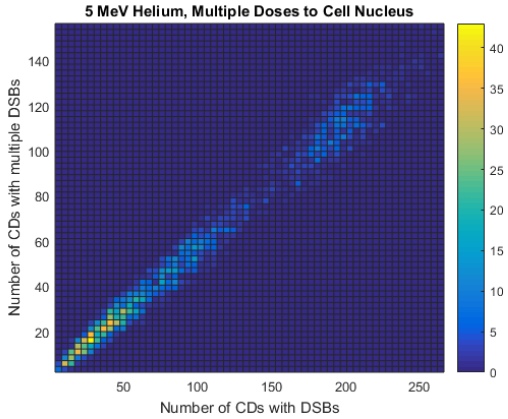
(a) 1 MeV Electrons



(b) 290 MeV/u Carbon Ions



(c) 17 MeV/u Carbon Ions



(d) 5 MeV Helium Ions

Figure 7.6: The number of CDs with multiple CF rods with DSBs vs. the number of CDs with one or more DSBs

x-axis values corresponding to the doses of 0.5 Gy and 1 Gy. The spread for these more sparsely ionizing types is also noted to be much more hazy, with relative ratio of CDs with multiple DSBs being far less defined, especially for the 1-MeV electrons.

7.2.3 Cell Survival Curves

As discussed in Section 7.1.3, the cell survival curves for various types of radiation can be produced by using the corresponding set of CFDs of DSB potential misrejoining events. As shown in Figures 7.3 and 7.5, each set of CFD corresponds to a specific radiation type and dose. To obtain the cell survival fraction for a specific radiation type and dose, one first applies a RNG to the CFD to obtain the corresponding number of potential misrejoining events. The cell is dead or alive is determined by specifying a cell killing probability for each potential misrejoining event and by applying a RNG to each of the potential misrejoining events: any RNG value above the specified cell killing probability for any one of the potential misrejoining events would be tallied as a cell killing event. Any cell not containing a value above the specified value was tallied as having survived the irradiation. This procedure is repeated for as many cells as needed until the result is converged to a preset statistical uncertainty. While the cell killing probability for each DSB misrejoining event can be different, it is assumed to be a constant which is adjusted to match the experimental cell survival curves. Figures 7.7 and 7.8 show the results of two sets of cell survival curves obtained from the above described procedure. The cell killing probability for each DSB misrejoining event was set to be 3%.

While basic, the results shown in Figures 7.7 and 7.8 bear a remarkable resemblance to the expected pattern for cell survival across multiple radiation types.

7.2.4 Hypoxia Simulations

An attempt is hereby presented for the simulation of hypoxia in our new model. One of the areas of great concern in radiotherapy is the impact of locally advanced tumors and

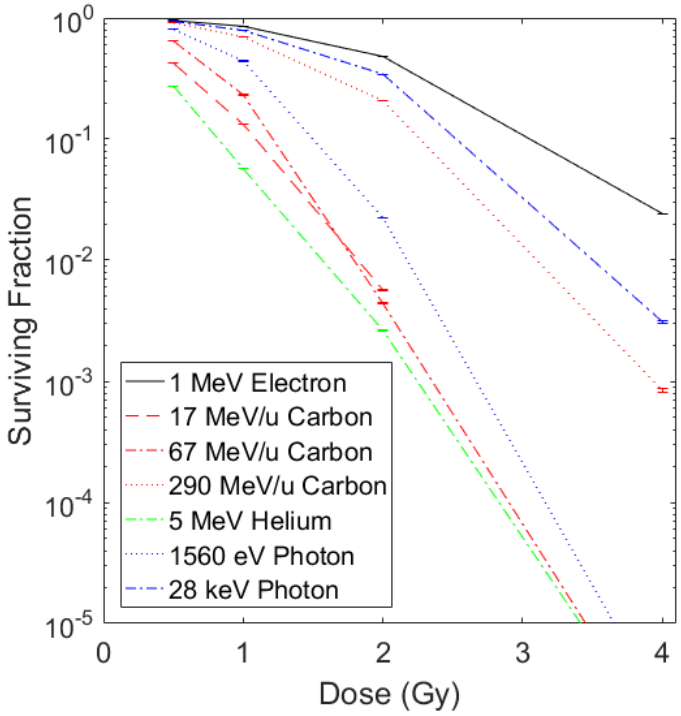


Figure 7.7: Cell survival curves obtained from this simulation study for various energies of electrons and protons. The cell killing probability for each potential DSB misrejoining event was set at 2%

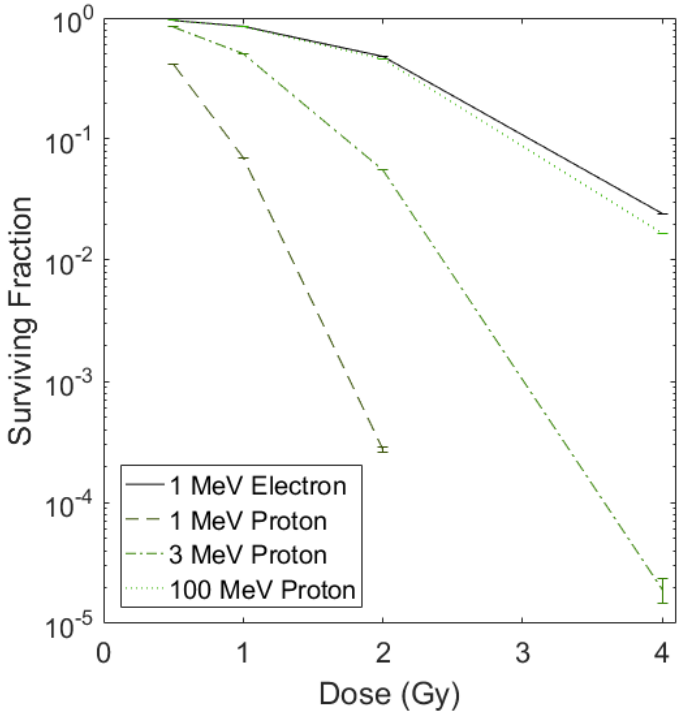


Figure 7.8: Cell survival approximation, 2% cell killing threshold for protons and electrons

the propensity for hypoxic cells farther from effective vasculature to have markedly higher resistance to the therapeutic effect of ionizing radiation treatment. It is one of the strong rationales in advocacy of high-LET radiotherapy, as higher LET particles have been noted to be less affected by the radioresistance induced by hypoxia. In the case of alpha particles, the literature regularly references an oxygen enhancement ratio (OER) of unity, indicating the cell survival of hypoxic cells matches the cell survival of normal cells in standard oxygenated conditions. For sparsely ionizing radiation types, the OER increases, being estimated as high as 2.5 for standard irradiation types such as cobalt-60. For intermediately dense ionizing radiation types the OER is somewhere in between, estimated as 1.5 in the case of a 28 MeV neutron beam.

The effect of hypoxia is mainly due to the lower availability of oxygen molecules in the cell nucleus to fixate DNA damage. As such we have put our focus on the theory that a reduced likelihood of DSB induction is the primary influence of hypoxia on cell survival calculations.

In our model, the simple way to simulate hypoxia is to increase the threshold energy for what can induce a DNA DSB. Our DSB calibration is based on a 40-50 DSB/Gy induction rate for sparsely ionizing radiation, which yielded a 75 eV threshold of energy cluster size for DSB induction. Several investigators in prior Monte Carlo simulations indicate this to have been a reasonable value. For hypoxia induction having a reduced availability of oxygen molecules to contribute to DNA damage, we simply increased the energy cluster size threshold required for a DNA DSB to 100 eV as an initial starting point. The results are seen in Table 7.1 regarding how much the DSB yield changes from our initial reference point.

Table 7.1: Comparison of the DSB yields of the oxic cells to that of the hypoxic cells obtained from the new model

	Normal			Hypoxic		
	DSBs/Gy	cDSBs/Gy	% cDSBs	DSBs/Gy	cDSBs/Gy	% cDSBs
1-MeV Electron	42.2	2.4	6%	23.7	1.0	4%
X-rays, 28 keV	50.8	3.0	6%	29.2	1.3	5%
X-rays, 1560 eV	83.7	6.5	8%	53.6	3.0	6%
X-rays, 286 eV	98.3	5.1	5%	65.2	2.3	4%
Protons, 100 MeV	44.7	2.6	6%	24.9	1.1	4%
Protons, 10 MeV	52.0	3.0	6%	29.3	1.3	4%
Protons, 3 MeV	67.5	4.1	6%	38.9	1.7	4%
Protons, 1 MeV	95.9	15.5	16%	71.8	8.6	12%
Helium, 400 MeV	44.9	2.6	6%	25.1	1.1	4%
Helium, 20 MeV	84.9	7.4	9%	55.0	3.5	6%
Helium, 5 MeV	65.2	33.8	52%	56.1	25.2	45%
Carbon, 290 MeV/u	55.0	3.1	6%	30.5	1.3	4%
Carbon, 185 MeV/u	61.1	3.5	6%	34.6	1.5	4%
Carbon, 67 MeV/u	82.8	6.4	8%	52.4	3.0	6%
Carbon, 17 MeV/u	69.1	19.2	28%	53.3	12.5	23%

As expected, the DSB yields decrease across all radiation types, with the differences being more pronounced for low-LET radiation types. A 40% reduction is seen for electrons,

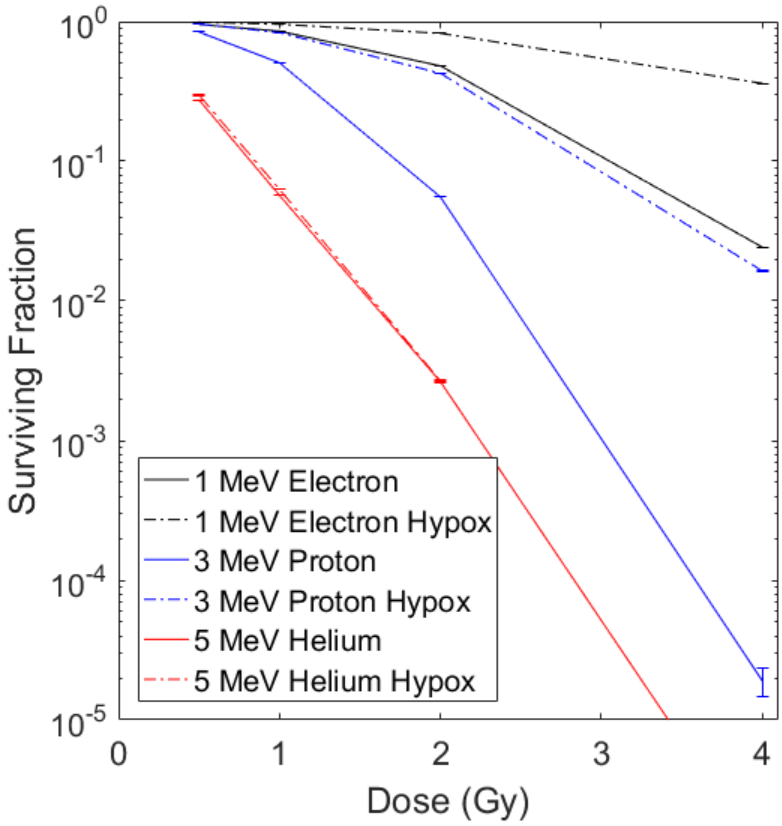


Figure 7.9: Cell survival approximation, hypoxia

28 keV photons, and 100 MeV protons, while a 14% reduction is seen for 5 MeV helium ions. In addition, the prevalence for cDSBs is seen to reduce in a similar fashion.

For cell survival curves, the simulations were re-run with the updated DSB data. It was expected that with the reduced number of DSBs and potential misrejoining events for all cell types, the simulation would not yield identical values for the reference 5 MeV helium ions as is observed in cell irradiation studies. In order to account for this, the cell killing probability for each potential DSB misrejoining event was increased from 2% to 3% so that cell survival curves under oxic and hypoxic conditions for 5 MeV helium ions become comparable. The results are seen in Figure 7.9.

For a matched cell killing of the 5 MeV helium ions, we can see by comparing Figure 7.9 to Figures 7.7 and 7.8 that the overall survival fraction of hypoxic cells of other radiation types are substantially higher than those of the oxic cells, which is expected given the

results in Table 7.1. With a more substantially reduced DSB yield relative to the reference conditions assumed from the high-LET helium ions, the low-LET radiation types have far fewer potential DSB misrejoining events in the simulations. Visual assessment of the ratios indicate an OER of 1 for 5 MeV helium ions, about 1.7 for the protons of both energies, and just around 2 for the 1-MeV electrons. While these results do not fully agree with that obtained from the experiment, they do follow the correct pattern.

7.2.5 Low-dose risk assessment for cancer

One of the obvious applications of this new model is in radiation protection, i.e. to provide low-dose risk assessment for cancer for various radiation types.. In radiation protection, low dose refers an effective dose that is less than 100 mSv [103]. Computational models become essential for low doses because the epidemiological data contain too big an uncertainty to be useful for risk assessment.

Currently, the radiation-induced cancer risks at low dose levels are estimated by linear extrapolation from the high-dose epidemiological data., which may be grossly incorrect. If one accepts the assumption that the initiation of cancer is largely due to exchange-type chromosome aberrations, which are mostly the result of DSB misrejoining, then the same set of CDFs used for calculating the cell survival curves can also be used to estimate the relative cancer risks among the various radiation types. In radiation protection, this relative risk is expressed by a quantity called the radiation weighting factor (W_R), which is defined in a similar way as that of RBE described in Section 2.2.1, i.e.

$$W_R = \frac{D_{X-ray}}{D_{test}} \tag{7.1}$$

where D_{test} is the dose of the test radiation in question to produce a specified cancer risk and $D_{X\text{-}ray}$ is the dose of X-ray required to produce the same cancer risk. Since by definition the W_R is estimated for low dose levels, we can simply use the CDFs of various radiation types at low dose levels (e.g. in the range between 0.1 cGy and 10 cGy) to obtain the W_R

for each radiation type.

The direct comparison of low-dose risk assessment from our model comes from a comparison of potential misrejoining event yields between different radiation types. Values comparing the yields between radiation types are shown in Table 7.2, with all columns normalized to the 1-MeV electron yields at 1 Gy dose levels. All values ignore interactions between tracks and assume single-track effects by individual particles.

Table 7.2: Relative proportions of different DSB and potential misrejoining types, single track effects normalized to 1 Gy

	cDSBs	Potential misrejoining events	CDs with multiple DSBs	Average of damage types
1- MeV Electrons	1.00	1.00	1.00	1.00
X-rays, 28 keV	0.99	1.81	1.81	1.05
X-rays, 1560 eV	1.23	0.88	1.76	1.21
X-rays, 286 eV	0.70			0.65
Protons, 100 MeV	1.08	2.49	4.97	1.21
Protons, 10 MeV	1.33	1.93	1.00	1.36
Protons, 3 MeV	1.35	3.29	2.83	1.49
Protons, 1 MeV	3.42	21.18	15.42	4.54
Helium, 400 MeV	0.97	1.41	1.97	1.01
Helium, 20 MeV	1.71	13.51	8.23	2.49
Helium, 5 MeV	8.28	21.29	19.13	8.93
Carbon, 290 MeV/u	1.21	2.42	2.02	1.28
Carbon, 185 MeV/u	1.30	3.31	2.32	1.42
Carbon, 67 MeV/u	1.82	11.96	8.08	2.49
Carbon, 17 MeV/u	5.51	31.73	22.47	6.90

Overall conclusions from Table 7.2 are preliminary, but the power of the model is demonstrated in that different forms of DNA damage likely have different contributions to cancer risk. The overall risk for alpha particles is partially demonstrated in the 5 MeV helium range, in that the starting energy of alpha decay is in this range. The W_R for alpha decay is taken as a value of 20, and Table 7.2 shows agreement with this number with some

measure of variability depending on how the DNA damage is assessed: the average of all damage types 14.15 is in the general range of what is expected, although a closer focus at the proportion of potential misrejoining of DSBs and possible micronucleus formation is quite a bit higher.

Results as shown in Table 7.2 indicate that our model has a very good application in radiation protection studies. As differences in damage can be compared in detail, the actual variation in danger from exposure can be more thoroughly explored.

CHAPTER 8
CONCLUSIONS

In the presented work, a new radiobiological model was developed entirely from principles of the Monte Carlo method. The model integrates charged particle track structures and the detailed cell nucleus DNA organizations to quantify DNA and chromatin damage. The DNA organizations in a cell nucleus include chromosome territories, chromatin domains, and chromatin fibers.

Several iterations were exhaustively explored for how to properly position chromatin fibers in the cell nucleus and how to analyze the charged particle track data in order to gain insight into spatially distributed radiation damage patterns. Several forms of radiation-induced DNA damage were evaluated, including sDSB, cDSB, DSB misrejoining, chromosome aberration, and cell death. The evaluation also includes varying dose levels of a broad range of radiation types. The evaluation demonstrates that the results obtained match well with that obtained experimentally over the broad range of radiation types. Moreover, we have shown that any form, dose, dose rate, and blend of different radiation types can be handled and evaluated using this new model.

Much of the literature from other radiobiological models hypothesizes potential causes for observed phenomena, but very few analyze the data in sufficient detail to explain how radiation damage is triggered. In this manner, our model offers a predictive element that we believe offers a powerful tool to experimental radiobiologists. Furthermore, our results indicate effects of low energy electron damage across a broad spectrum of radiation types in a way that has been expected, but not yet quantified in other models. The statistical breakdown presented in our model has also been useful in producing preliminary cell survival curves based on a cell-by-cell simulation, which is absent in the most prevalent analytical models.

Appendices

APPENDIX A
REVISIONS TO EARLIER WORK

A.1 Kinetic energy of particles causing DSBs

Given the results in Table 6.3 showing individual DSBs are the result of single-track effect, a closer look into how these single tracks cause DSBs was considered beneficial. For each cluster of energy depositions leading to DSBs, the KE of the particle for each event is output. If a single particle causes 10 events, the instantaneous KE of the particle is reported with each event and provides some information on the nature of the involved particles in the cluster. This analysis provides information on the particles directly involved in causing DSBs.

When an energetic particle produces a very low energy delta ray, the majority of clustered energy deposition is due to the delta ray and not the more energetic particle. In other words, there is some benefit to ignoring the highest energy particle in a clustered energy deposition if the high energy particle is only contributing to the DSB to a minute extent. Table A.1 shows data on clusters excluding the energy deposition event with the highest KE. What is meant is that the kinetic energy of the responsible particle for every individual energy deposition is tabulated, and only the single highest event is excluded: if a high energy particle causes two ionization events, it would still be included, but if it only contributes a single time it would be filtered out.

While the energy of electrons and delta rays can be fairly high (greater than 5 keV, sometimes substantially so), Table A.1 shows the proportion of clusters identified by DBSCAN wherein the highest energy involved particle is below 1 keV is 89% for a reference high-energy electron spectrum. This indicates that the bulk of damage leading to DSBs is provided by charged particles of low enough KE so as to cluster their damages more

Table A.1: Second highest particle kinetic energy value in a DSB

	<1 keV	1 keV - 100 keV	>100 keV
Electrons	89%	10%	1%
X-rays, 28 keV	85%	15%	0%
Protons, 100 MeV	89%	8%	3%
Protons, 10 MeV	74%	8%	18%
Protons, 3 MeV	51%	5%	44%
Protons, 1 MeV	17%	0%	83%
Helium, 400 MeV	82%	8%	10%
Helium, 20 MeV	29%	4%	67%
Helium, 5 MeV	19%	0%	81%
Carbon, 290 MeV/amu	55%	7%	38%
Carbon, 185 MeV/amu	47%	6%	47%
Carbon, 67 MeV/amu	31%	5%	64%
Carbon, 17 MeV/amu	30%	5%	65%

severely. This scenario is similarly seen for high energy protons and helium ions, indicating their delta rays are the primary direct source of clustered energy depositions. For high LET particles, and carbon ions in general, a more substantial contributor of energy depositions comes from the heavy ion itself, indicating the nature of DSB production by high-LET particles is due to the higher spatially connected behavior of the primary particle itself.

A.2 Revisions to earlier work

Early validation of the technical approach focused on confirming the model was consistent with expectation values of energy deposition to the target structures and to the cell nucleus. Literature reviews have indicated one of the challenges for cell and small animal irradiation experiments is inconsistent dosimetry, with variations from expectation values sometimes exceeding 40%. We wanted to make sure, as much as possible, the values were consistent.

As the goal of the presented work is to develop a consistent Monte Carlo model for radiobiology, the initial proportion of results aimed to validate the inputs and setup parameters for the model. This consisted of the confirmation of dose consistency, determination of clustering parameters, and modeling of DSB production. Details regarding the model were also analyzed to see if valid observations regarding the origination of DSBs could be assessed from our model.

The later portion of results breakdown focused on how to extract viable data from the model that would be pertinent for cell irradiation studies. This includes assessment of the proximity of DSBs so as to determine the likelihood of different quantifiable values, including cell survival, micronucleus production, and the like. This portion is a continuing effort, the results of which are more pertinent for direct application.

A.2.1 Migration from a pure MFP approach to comprehensively simulating all structures in the cell nucleus

With respect to the earliest assessments, model validation takes a simple form: the energy deposited within the cell simply needs to be consistent. It is understood that chromatin fiber occupies 4% of the total volume of the cell nucleus, and as such the total energy deposition to chromatin fiber structures similarly should be close to 4% for all studies, irrespective of the gross quantity or radiation type. Large variations are possible regarding the quantity of clustered radiation damage, the total step lengths traversed by particles in their paths, and their typical patterns, but the overarching consistent value should be that the total energy deposited in the cell nucleus should be proportional to the total dose, and the proportion of energy deposited within the CF rods will be consistent across radiation types.

This step was critically important in the decision to migrate from a pure step-length-based mean-free-path for modeling CF rods to the application of fully modeling the CF rods within a cell nucleus beforehand. The mean-free-path approach was initially adopted to be more computationally efficient: by seeding CF rods only along a particle path, the

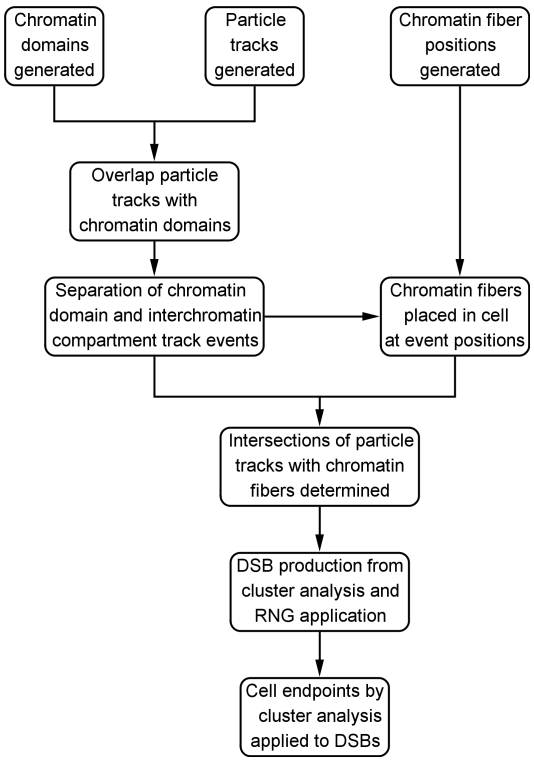
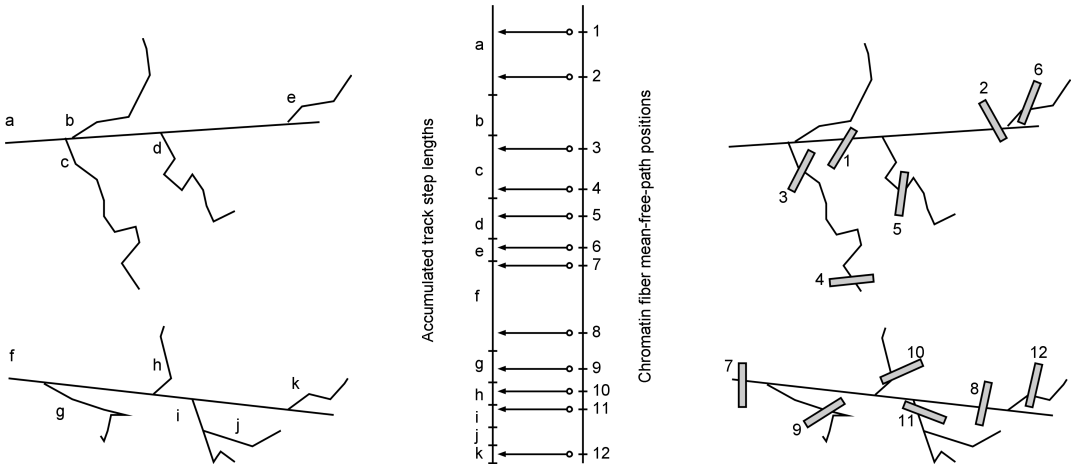


Figure A.1: Obsolete flowchart for model using a pure mean-free-path approach

total number of detection regions could be reduced by orders of magnitude for most simulations. This approach was sufficient for more sparsely ionizing radiation types, such as high energy electrons, restricted photon simulations (exclusion of the photon step lengths and only looking at the directly ionizing electron depositions), and high energy heavier ions. However, for high-LET particles this method overestimated the total energy deposition to CF rods. Table (insert table here) shows the pure MFP approach led to a total energy deposition to CF rods of as high as 9% for 5 MeV helium ions, more than a factor of 2 in excess of what should be reported.

A modified MFP approach using a grid of rays was used to produce CF rods and the cells were stored separately. These cell nucleus models were subsequently used in the simulations and the total energy deposition to the CF rods match the 4% value very closely. Substantial efficiency improvements to the intersection algorithm over the course of this work contributed to the viability of the comprehensive approach, as well. Further cell



(a) Linear grid view for MFP placement of CF rods

(b) Linear grid view for MFP placement of CF rods

(c) Ladder method for determining MFP placement of CF rods

Figure A.2: MFP placement of CF rods along fixed grids with stochastic seeding

simulations were performed by making use of the generated cell models.

Further Notes

The decision to abandon the MFP approach based on following charged particle tracks through the cell nucleus was challenging: over a full year of effort of the past two years was dedicated to perfecting this novel approach. However, we were able to translate the core principles into the production of cell nucleus templates in the grid of rays application. Moreover, it provided the initial stepping stone for migrating from a single-track simulation to a fully-translated simulation of a cell nucleus in spatial coordinates.

A.3 Codes

A.3.1 Intersection Code

```
function []=clusterCT(PNRG,Pfrac,Ptype,series)
parfor i=1:96
    Ptypes=['A','C','E','P','X'];
    fileID=fopen(strcat('cell.',num2str(i+96*(series-1)),'.bin'));
    CFs=fread(fileID,'double');  fclose(fileID);  CFs=reshape(CFs,[],13);
```

```
ICCFs=CFs(CFs(:,12)==0,:);  CDCFs=CFs(CFs(:,12)~=0,:);
ii=CFs(1,3);  CFs=[];
fileID=fopen(strcat(Ptypes(Ptype),num2str(PNRG),'.D',num2str(Pfrac),'dat.',int2str(ii)
    ,'.bin'));
primmy=fread(fileID,'single');  fclose(fileID);  primmy=reshape(primmy,[],12);
Points=zeros(18,3); evtqtyIC=0; evtqtyCD=0;
ICCFa=ICCFs(:,4:9); ICCFb=ICCFa(:,4:6)-ICCFa(:,1:3);
ICCFa(:,1:3)=ICCFa(:,1:3)+0.25*ICCFb(:,1:3);
ICCFa(:,4:6)=ICCFa(:,4:6)-0.25*ICCFb(:,1:3);
CDCFa=CDCFs(:,4:9); CDCFb=CDCFa(:,4:6)-CDCFa(:,1:3);
CDCFa(:,1:3)=CDCFa(:,1:3)+0.25*CDCFb(:,1:3);
CDCFa(:,4:6)=CDCFa(:,4:6)-0.25*CDCFb(:,1:3);
[~,ICDa] = knnsearch(primmy(:,6:8),ICCFa(:,1:3));
[~,ICDb] = knnsearch(primmy(:,6:8),ICCFa(:,4:6));
IChita=ICDa<43; IChitb=ICDb<43;
IChitc=IChita+IChitb;   IChitc=IChitc~=0;
[~,CDDa] = knnsearch(primmy(:,6:8),CDCFa(:,1:3));
[~,CDDb] = knnsearch(primmy(:,6:8),CDCFa(:,4:6));
CDhita=CDDa<43; CDhitb=CDDb<43;
CDhitc=CDhita+CDhitb;   CDhitc=CDhitc~=0;
ICCFs=ICCFs(IChitc,:);  CDCFs=CDCFs(CDhitc,:);
[dIC1,ICD1] = knnsearch(ICCFa(IChitc,1:3),primmy(:,6:8));
[~,ICD2] = knnsearch(ICCFa(IChitc,4:6),primmy(:,6:8));
IChit1=ICD1<43; IChit2=ICD2<43;
IChit3=IChit1+IChit2;   IChit3=IChit3~=0;
IChitb=[];  CDhitb=[];  IChita=[];  CDhita=[];  CDDa=[];  CDDb=[];  ICDa=[];  ICDb=[];
[dCD1,CDD1] = knnsearch(CDCFa(CDhitc,1:3),primmy(:,6:8));
[~,CDD2] = knnsearch(CDCFa(CDhitc,4:6),primmy(:,6:8));
CDhit1=CDD1<43; CDhit2=CDD2<43;
CDhit3=CDhit1+CDhit2;   CDhit3=CDhit3~=0;
CDhits=[primmy(CDhit3,:) CDCFs(dCD1(CDhit3,:),12:13)];
IChits=[primmy(IChit3,:) ICCFs(dIC1(IChit3,:),12:13)];
CDhitsize=size(CDhits,1);   IChitsize=size(IChits,1);
ICchrom=size(ICCFs,1);  CDchrom=size(CDCFs,1);
primmy=[]; CDhit1=[];  CDhit2=[];  CDhit3=[];  IChit1=[];  IChit2=[];  IChit3=[];
    CDD1=[];    CDD2=[];  dCD1=[];  dIC1=[];  ICD1=[];  ICD2=[];  ICCFa=[];  CDCFa=[];
      IChitc=[];  CDhitc=[];
if isempty(CDCFs);
else
    inIC=zeros(IChitsize,1);  inICP=inIC;
    for chromIC=1:ICchrom
```

```
            IChitIDX=IChits(:,13)==ICCFs(chromIC,12);
            Points = cylinder2Pc(15,ICCFs(chromIC,4:6),ICCFs(chromIC,7:9));
            inICP(IChitIDX)=inpolyhedron(freeBoundary(delaunayTriangulation(Points)),
                Points,IChits(IChitIDX,6:8));
            eventqty=sum(inICP(IChitIDX));
            if eventqty==0
            else
                evtqtyIC=evtqtyIC+1;
                try
                    inIC(inIC(:)==0)=bsxfun(@times,inICP(inIC(:)==0),ICCFs(chromIC,11));
                catch
                end
            end
        end
        inCD=zeros(CDhitsize,1);  inCDP=inCD;
        for chromCD=1:CDchrom
            CDhitIDX=CDhits(:,13)==CDCFs(chromCD,12);
            Points = cylinder2Pc(15,CDCFs(chromCD,4:6),CDCFs(chromCD,7:9));
            inCDP(CDhitIDX)=inpolyhedron(freeBoundary(delaunayTriangulation(Points)),
                Points,CDhits(CDhitIDX,6:8));
            eventqty=sum(inCDP(CDhitIDX));
            if eventqty==0
            else
                evtqtyCD=evtqtyCD+1;
                try
                    inCD(inCD(:)==0)=bsxfun(@times,inCDP(inCD(:)==0),CDCFs(chromCD,11));
                catch
                end
            end
        end
        if (sum(inCD))==0
            events=[];
        else
            events=[IChits,inIC;CDhits,inCD];
        end
        fileID=fopen(strcat(Ptypes(Ptype),num2str(PNRG),'D',num2str(Pfrac),'.',int2str
            (96*(series-1)+i),'.bin'),'w');
        fwrite(fileID,events(events(:,15)~=0,:),'double');  fclose(fileID);
    end
end
end
```

A.3.2 Clustering Code

```
function []=scannerCT(PNRG,Pfrac,Ptype,series)
clustsize=2;  radius=[2,4,1,3];  Ptypes=['A','C','E','P','X'];
for radiu=1
    parfor proc=1:96
        cluma=zeros(1000000,18);  groggy=1;  pror=(series-1);  runs=96*pror+proc;
        try
            fileID=fopen(strcat(Ptypes(Ptype),num2str(PNRG),'D',num2str(Pfrac),'.',int2str
                (runs),'.bin'));
            chroms=fread(fileID,'double');  fclose(fileID);  chroms=reshape(chroms,[],15);
        catch
            chroms=[];
        end
        if isempty(chroms)
        else
            bema=unique(chroms(:,13));  fibmax=unique(chroms(:,15));
            for fibid=fibmax'
                chrma=chroms(chroms(:,15)==fibid,:);  clug=chrma(:,6:8);
                try
                    [clui,idtype]=dbscan1(clug,clustsize,radius(radiu));
                catch
                    clui=[];  idtype=[];
                end
                if isempty(clui)
                else
                    idtype=idtype';  clui=clui';
                    clda=[chrma((idtype(:)~=(-1)),[6:9,5,13,1]),clui(idtype(:)~=(-1)),
                        chrma((idtype(:)~=(-1)),14)];
                    for beid=bema'
                        befi=clda(clda(:,6)==beid,:);
                        if isempty(befi)
                        else
                            clma=unique(befi(:,8));
                            for clbe=clma'
                                bfc=befi(befi(:,8)==clbe,:);
                                elii=[bfc(bfc(:,5)==13,2);bfc(bfc(:,5)==18,2);bfc(bfc(:,5)
                                    ==24,2);bfc(bfc(:,5)==35,2)];
                                ionqty=size(elii,1);
                                elei=[bfc(bfc(:,5)==12,2);bfc(bfc(:,5)==17,2);bfc(bfc(:,5)
                                    ==23,2);bfc(bfc(:,5)==36,2)];
```

```
                                exciqty=size(elei,1);  ioexqty=ionqty+exciqty;
                                ionrg=sum(bfc(:,4));  iosz=size(bfc,1);
                                clupo=[(sum(bfc(:,1))/iosz),(sum(bfc(:,2))/iosz),(sum(bfc
                                    (:,3))/iosz)];
                                trackE=unique(bfc(:,7));  trun=zeros(1,5);
                                if size(trackE,1)>5
                                    trun=trackE(1:5)';
                                else
                                    for etrack=1:size(trackE,1)
                                        trun(etrack)=trackE(etrack);
                                    end
                                end
                                clusu=[runs,beid,fibid,clbe,ionqty,exciqty,ioexqty,ionrg,
                                    clupo,trun,size(trackE,1),bfc(1,9)];
                                try
                                    cluma(groggy,:)=clusu  groggy=groggy+1;
                                catch err
                                    cluma=[cluma;zeros(1000000,18)];
                                    cluma(groggy,:)=clusu;  groggy=groggy+1;
                                end
                            end
                        end
                    end
                end
            end
        end
        cluma(~any(cluma,2),:)=[];
        fileID=fopen(strcat(Ptypes(Ptype),num2str(PNRG),'D',num2str(Pfrac),'clust',int2str
            (runs),'C',int2str(clustsize),'R',num2str(radius(radiu)),'.bin'),'w');
        fwrite(fileID,cluma,'double');  fclose(fileID);
    end
end
end
```

A.3.3 DSB RNG code

```
% Runs RNG through raw cluster data from combined scannerCT data for DSBs.
% Probability of DSBs based on energy deposited to clusters. Method uses
% 75 eV as a DSB threshold, for each 75 eV a cluster contains, a uniform
% RNG is run to determine probability of DSB from cluster.  RNG > 0.90 is
% scored as a DSB.  cDSB where multiple RNG values of a cluster > 0.90.
```

```
% Rationale for 0.90 was a calibration to bring the DSB production of 1 Gy
% electron radiation at 40 DSB/Gy.
partname=['E0.5D0.5 ';'E0.5D1   ';'E0.5D2   ';'E0.5D4   ';..;'A400D10  '];
radius=[2,1,3,4];
for radiu=1
 yieldDSB=[];
    parfor partnum=1:64
        fileID=fopen(strcat(strtrim(partname(partnum,:)),'clustSR',num2str(radius(radiu)),
            '.bin'));
        clusterdata=fread(fileID,'double');  fclose(fileID);
        clusterdata=reshape(clusterdata,[],18);
        [~,idc,~]=unique(clusterdata(:,1:3),'rows');  ide35=clusterdata(idc,:);
        [~,idx,ic]=unique(clusterdata(:,1:4),'rows');  datE1=clusterdata(idx,:);
        idx35E1=find(datE1(:,8)>75,size(datE1,1));  dtE1=datE1(idx35E1,:);
        iE1=int8(dtE1(:,8)/75);
        rngE1=zeros(size(iE1,1),max(iE1));
        for cb=1:size(iE1,1)
            rngE1(cb,1:iE1(cb))=rand(1,iE1(cb));
        end
        rnghitE1=sum((rngE1>0.900),2);  rnghit1=rnghitE1==1;  rnghit2=rnghitE1==2;
            rnghit3=rnghitE1==3;  rnghit4=rnghitE1>3;
        dtE1a=[dtE1 rnghitE1];
        dtE1b=dtE1a(dtE1a(:,19)~=0,:);  dtE1b=sortrows(dtE1b,[1 2 3 -19]);
        [~,idt1,idtx]=unique(dtE1b(:,1:3),'rows');
        dtE1c=dtE1b(idt1,:);
        idt2=diff(idt1);
        idt2=[idt2;size(dtE1b(dtE1b(:,3)==dtE1b(size(idt2,1),3),:),1)];
        dtE1b(:,16)=idt2(idtx);
        szDt1=unique(idt1(:));
        for szD=1:size(szDt1,1)
            clusterdata(clusterdata(:,3)==szDt1(szD),16)=idt2(idt1==szDt1(szD));
        end
        idxhit1=find(rnghit1,size(rnghitE1,1));
        iOut1=ismember(ic,idx35E1(idxhit1));  out1=clusterdata(iOut1,:);
        out1=[out1(:,1:15) ones(size(out1,1),1) out1(:,17:18)];
        idxhit2=find(rnghit2,size(rnghitE1,1));
        iOut2=ismember(ic,idx35E1(idxhit2));  out2=clusterdata(iOut2,:);
        out2=[out2(:,1:15) ones(size(out2,1),1)*2 out2(:,17:18)];
        idxhit3=find(rnghit3,size(rnghitE1,1));
        iOut3=ismember(ic,idx35E1(idxhit3));  out3=clusterdata(iOut3,:);
        out3=[out3(:,1:15) ones(size(out3,1),1)*3 out3(:,17:18)];
```

```
        idxhit4=find(rnghit4,size(rnghitE1,1));
        iOut4=ismember(ic,idx35E1(idxhit4));  out4=clusterdata(iOut4,:);
        out4=[out4(:,1:15) ones(size(out4,1),1)*4 out4(:,17:18)];
        outE1=[out1;out2;out3;out4];  outE1=sortrows(outE1,[1 2 3 4]);
        fileID=fopen(strcat(strtrim(partname(partnum,:)),'clustDR',num2str(radius(radiu)),
            '.bin'),'w');
        fwrite(fileID,outE1,'double');  fclose(fileID);
        cellno=max(datE1(:,1));
        dsbE1=size(unique(outE1(:,1:3),'rows'),1)/cellno;
        datBE1=size(unique([out2(:,1:3);out3(:,1:3);out4(:,1:3)],'rows'),1);
        yieldDSBa=[partnum,size(ide35,1)/cellno,size(idx35E1,1)/cellno,size(dtE1c,1)/
            cellno,size(dtE1c(dtE1c(:,19)>1,:),1)/cellno,size(idt2(idt2>1,:),1)/cellno,0,
            sum(datE1(:,8))/cellno,sum(dtE1(:,8))/cellno,sum(dtE1b(:,8))/cellno,0,sum(dtE1
            (dtE1(:,8)>150,8))/cellno];
        yieldDSB=[yieldDSB;yieldDSBa];
    end
    yieldDSB=sortrows(yieldDSB,1);
    csvwrite(strcat('DSBYield',num2str(radius(radiu)),'.csv'),yieldDSB);
end
```

A.3.4 DSB Proximity Code

```
function [result] = ClusterProxPCD4a(hitsize,raddsb)
% Opens the DSB yield from DSBconvert.m, runs DBSCAN on DSBs with input
% parameters hitsize and raddsb.  Outputs cluster data of DSBs based on
% proximity to DSBs, also tags unclustered DSBs
ptna=['E0.5D0.5 ';'E0.5D1   ';'E0.5D2   ';'E0.5D4   ';..;'A400D10  '];
radi=2;  result=[];
    parfor ptnu=1:64
        try
            fileID=fopen(strcat(strtrim(ptna(ptnu,:)),'clustDR',num2str(radi),'.bin'));
            clusterdata=fread(fileID,'double');  fclose(fileID);
            clusterdata=reshape(clusterdata,[],18);
            data2=clusterdata(clusterdata(:,8)>hitsize,:);
            cellno=unique(data2(:,1));
            for cellcounter=1:(size(cellno,1))
                cellcount=cellno(cellcounter);
                data3=data2(data2(:,1)==cellcount,:);  clustdsb=1;
                [clustid,idtype]=dbscan1(data3(:,9:11),clustdsb,raddsb);
                idtype=idtype';  clustid=clustid';
                dsbdata=[data3((idtype(:)~=(-1)),:),clustid(idtype(:)~=(-1))];
```

```
                [~,IA,~]=unique(dsbdata(:,[1:3,17:18]),'rows');
                dsbdatatest=dsbdata(IA,:);  dsbdatatest=sortrows(dsbdatatest,19);
                sv=dsbdatatest(:,18);
                IB = sv(2:end) == sv(1:end-1);
                res = dsbdatatest(ismember(dsbdatatest(:,19),sv(IB)),:);
                resn= dsbdatatest(~ismember(dsbdatatest(:,19),res(:,19)),:);
                dsbig=[data3((idtype(:)==(-1)),:),clustid(idtype(:)==(-1))];
                [~,IBig,~]=unique(dsbig(:,[1:3,17:18]),'rows');
                dsbig=dsbig(IBig,:);  resn=[resn;dsbig];  resn(:,19)=0;
                res=[res;resn];  res=[res (ones(size(res,1),1))*ptnu];
                result=[result;res];
            end
        catch
        end
    end
    fileID=fopen(strcat('cellclusterStDec',num2str(radi),'.H.',num2str(hitsize),'.D.',
        num2str(raddsb),'.bin'),'w');
    fwrite(fileID,result,'double');  fclose(fileID);
end
```

A.3.5 Potential DSB Misrejoining Cluster Combinations Code

```
function [cctr] = ClustPrxB1(htz,rds)
% This script opens dt from  ClusterProxPCD4a.m.
radi=2;
fileID=fopen(strcat('cellclusterStDec',num2str(radi),'.H.',num2str(htz),'.D.',num2str(rds)
    ,'.bin'));
rlt=fread(fileID,'double');  fclose(fileID);  rlt=reshape(rlt,[],20);
%     Particle Type - Cell Number - CD Number - CF ID - Cluster ID - Energy - Tracks in
    Clust - Track1 - Track2
reu=[rlt(:,20) rlt(:,1) rlt(:,2) rlt(:,3) rlt(:,19) rlt(:,8) rlt(:,17) rlt(:,12) rlt(:,13)
     rlt(:,16) rlt(:,18)];
clearvars rlt
cctr=zeros(size(reu,1),15);  cctrid=1;
for partnum=1:64
        dt2=reu(reu(:,1)==partnum,:); % organize by particle type
        cellno=unique(dt2(:,2));
        for cecter=1:(size(cellno,1)) % organize by cell number
            cect=cellno(cecter);  res=dt2(dt2(:,2)==cect,:);
            vs=unique(res(:,5));
                for clnoct=1:size(vs,1)  % select by cluster number
```

```
                    clno=vs(clnoct);
                    if clno<1
                        dt4=res(res(:,5)==(clno),:);  cct=zeros(1,15);
                        cct(1:8)=[partnum cect clno size(dt4(dt4(dt4(:,3)==0,10)>1,:),1)
                            size(dt4(dt4(dt4(:,3)~=0,10)>1,:),1) size(dt4(dt4(:,3)==0,:)
                            ,1) size(dt4(dt4(:,3)~=0,:),1) -1];
                    else
                        dt4=res(res(:,5)==(clno),:);
                        dt5=dt4(dt4(:,3)~=0,:);  dt5=sortrows(dt5,3);
                        cct=zeros(1,15);  mtcont=size(unique(dt5(:,8)),1);
                        ccCDct=unique(dt5(:,3));  ccCDid=size(ccCDct,1);
                        ccCDiqt=zeros(ccCDid,1);
                        for ccCDcount=1:ccCDid
                            ccCDiqt(ccCDcount)=size(dt5(dt5(:,3)==ccCDct(ccCDcount),:),1);
                        end
                        ccCDiqt=(sortrows(ccCDiqt,[-1]))';
                        if ccCDid>3
                            cct(9:11)=ccCDiqt(1:3);
                        else
                            cct(9:(ccCDid+8))=ccCDiqt(1:ccCDid);
                        end
                        cct(13)=size(ccCDiqt(ccCDiqt(:)>1),2);
                        cct(14)=size(unique(dt4(dt4(:,11)~=0,11)),1);
                        cct(15)=size(unique(dt4(:,8)),1);
                        dt6=sortrows(dt4(dt4(:,11)~=0,11));
                        [~,ia,~]=unique(dt6);  ia=[ia;size(dt6,1)];
                            cct(1:8)=[partnum cect clno size(dt4(dt4(dt4(:,3)==0,10)>1,:)
                                ,1) size(dt4(dt4(dt4(:,3)~=0,10)>1,:),1) size(dt4(dt4(:,3)
                                ==0,:),1) size(dt4(dt4(:,3)~=0,:),1) ccCDid];
                            cct(12)=cct(6)+cct(8);
                            if cct(12)==2
                                cct(12)=1; % sets potential misrejoining of 2 DSBs to 1
                            end
                    end
                cctr(cctrid,:)=cct;  cctrid=cctrid+1;
                end
        end
cctr=cctr(any(cctr,2),:);
csvwrite(strcat('clusteroutputB1A.',num2str(radi),'.H.',num2str(htz),'.D.',num2str(rds),'.
    csv'),cctr);
end
```

A.3.6 Potential DSB Misrejoining Per Cell Code

```
function [cz,ceav] = ClustPrxB2(htz,rds)
% This script takes outputs of ClustPrxB1 and reduces it to values per cell
r=2;  ceav=zeros(64,9);  cz=zeros(1000000,7);  iccl=1;
result=csvread(strcat('clusteroutputB1A.',num2str(r),'.H.',num2str(htz),'.D.',num2str(rds)
    ,'.csv'));
for panu=1:64
    try
        da2=result(result(:,1)==panu,:);
        ceno=unique(da2(:,2));
        for ccst=1:(size(ceno,1))
            ceco=ceno(ccst);  dace=da2(da2(:,2)==ceco,:);
            noncl=dace(dace(:,3)==0,:); % Only data not in clusters
            dacl=dace(dace(:,3)>0,:);  % Only data within clusters
            if isempty(noncl)
                if isempty(dacl)
                    cz(iccl,1:2)=[panu ceco 0 0 0 0 0];  iccl=iccl+1;
                else
                    cz(iccl,:)=[panu ceco max(dacl(:,3)) 0 0 sum(dacl(:,12)) sum(dacl
                        (:,13))];
                    iccl=iccl+1;
                end
            elseif isempty(dacl)
                cz(iccl,:)=[panu ceco 0 noncl(4) noncl(5) 0 0];  iccl=iccl+1;
            else
                cz(iccl,:)=[panu ceco max(dacl(:,3)) noncl(4) noncl(5) sum(dacl(:,12)) sum
                    (dacl(:,13))];
                iccl=iccl+1;
            end
        end
        ceav(panu,:)=[panu mean(cz(cz(:,1)==panu,3:7),1) sum(mean(cz(cz(:,1)==panu,3:7),1)
            ), std(sum(cz(cz(:,1)==panu,3:7),2)) (sum(mean(cz(cz(:,1)==panu,3:7),1))-2*std
            (sum(cz(cz(:,1)==panu,3:7),2)))];
    catch
    end
end
cz=cz(cz(:,1)~=0,:);
csvwrite(strcat('clstprxB2.',num2str(r),'.H.',num2str(htz),'.D.',num2str(rds),'.csv'),cz);
csvwrite(strcat('clstprxB2a.',num2str(r),'.H.',num2str(htz),'.D.',num2str(rds),'.csv'),
    ceav);
```

A.3.7 Cell Survival Code

```
function[dz]=SurvB(htz,rds)
dz=zeros(99,64);  partname=['E0.5D0.5 ';..;'X28D4    '];  r=2;
rz=csvread(strcat('clstprxB2.',num2str(r),'.H.',num2str(htz),'.D.',num2str(rds),'.csv'));
parfor ii=1:64
    try
        data=rz(rz(:,1)==ii,:);  hitc=zeros(1000000,4);
        data(:,3)=sum(data(:,4:7),2);
        for cl=4:7
            [fPDF,vPDF]=ecdf(data(:,3));
            rndv=rand(1000000,1);  intv=zeros(size(rndv,1),1);
            for cellc=1:size(rndv,1)
                try
                  intv(cellc)=max(vPDF(fPDF<rndv(cellc)));
                catch
                end
            end
            hitv=zeros(size(rndv,1),max(intv));
            for cellc=1:size(rndv,1)
                hitv(cellc,1:intv(cellc))=rand(1,intv(cellc));
            end
            hitc(:,(cl-3))=max(hitv,[],2);
        end
        for thresh=1:199
            rsrv=zeros(size(hitv,1),1);  rkll=zeros(size(hitv,1),1);
            for cellc=1:size(hitv,1)
                if any(hitc(cellc,1:4)>(thresh/200))==0
                    rsrv(cellc)=1;
                else
                    rkll(cellc)=1;
                end
            end
            dz(thresh,ii)=sum(rsrv);
        end
    catch
    end
    fileID=fopen(strcat('survB.R',num2str(r),'.D',num2str(rds),'.',num2str(ii),'.bin'),'w'
        );  fwrite(fileID,hitc,'double');  fclose(fileID);
end
csvwrite(strcat('SurvB.',num2str(r),'.H.',num2str(htz),'.D.',num2str(rds),'.csv'),dz);
```

REFERENCES

[1] E Hall and A Giaccia, *Radiobiology for the radiobiologist*, Sixth. Philadelphia: Lippincott Williams & Wilkins, 2006.

[2] D Goodhead, *An assessment of the role of microdosimetry in radiobiology*, ser. 45-76 1. Radiation Research Society Res, 1982, vol. 91.

[3] D Charlton, D Goodhead, W Wilson, and H Paretzke, "Energy deposition in cylindrical volumes," M.R.C. Radiobiology Unit, Chilton, UK, Monograph 85/1, 1985.

[4] L Carter and E Cashwell, "Particle-transport simulation with the monte carlo method," Los Alamos Scientific Laboratory, Tech. Rep. TID-26607, 1975.

[5] S Incerti and et al., "The geant4-dna project," *Int j model simul sci comput*, vol. 01, pp. 157–178, 2010.

[6] S Incerti, A Ivanchenko, and et al., "Comparison of geant4 very low energy cross section models with experimental data in water," *Med phys*, vol. 37, no. 9, pp. 4692–4708, 2010.

[7] W Friedland, M Dingfelder, P Kundrát, and P Jacob, "Track structures, dna targets and radiation effects in the biophysical monte carlo simulation code partrac," *Mutat res*, vol. 711, pp. 28–40, 2011.

[8] A Ponomarev, J Huff, and F Cucinotta, "The analysis of the densely populated patterns of radiation-induced foci by a stochastic, monte carlo model of dna double-strand breaks induction by heavy ions," *Int j radiat biol*, vol. 86, no. 6, pp. 507–515, 2010.

[9] J Fowler, "Development of radiobiology for oncology - a personal view," *Phys med biol*, vol. 61, R263–R286, 2006.

[10] P Mayles, A Nayhum, and J Rosenwald, *Handbook of radiotherapy physics theory and practice*. Tayler and Francis Group, LLC, 2007.

[11] F Yin, J Wong, and et al., "The role of in-room kv x-ray imaging for patient setup and target localization," *Med phys*, 2009.

[12] M Herman, J Balter, and et al., "Clinical use of electronic portal imaging: Report of aapm radiation therapy committee task group 58," *Med phys*, vol. 28, no. 5, 2001.

[13] N Papanikolaou, J Battista, and et al., "Tissue inhomogeneity corrections for megavoltage photon beams," American Association of Physicists in Medicine, Tech. Rep., 2004.

[14] *Intensity-modulated radiation therapy: The state of the art*. American Association of Physicists in Medicine, 2003.

[15] J Palta and T Mackie, *Uncertainties in external beam radiation therapy*. American Association of Physicists in Medicine, 2011.

[16] E Klein, J Hanley, and et al., "Task group 142 report: Quality assurance of medical accelerators," *Med phys*, vol. 36, no. 9, 2009.

[17] M. D. Santos, C Villagrasa, I Clairand, and S Incerti, "Influence of the dna density on the number of clustered damages created by protons of different energies," *Nucl instr & methods in phys res b*, vol. 298, pp. 47–54, 2013.

[18] B Lee, "A monte carlo investigation of radiation damage to chromatin fibers and production of dna double strand breaks using geant4-dna code," Master's thesis, Georgia Institute of Technology, Atlanta, GA, 2014.

[19] J Lu and L Brady, *Radiation oncology: An evidence-based approach*. Springer-Verlag Berlin Heidelberg, 2008.

[20] E Hansen and M Roach, *Handbook of evidence-based radiation oncology*. Springer, 2010.

[21] H Albiez, M Cremer, and et al., "Chromatin domains and the interchromatin compartment form structurally defined and functionally interacting nuclear networks," *Chromosome research*, vol. 14, pp. 707–733, 2006.

[22] M Branco and A Pombo, "Chromosome organization: New facts, new models," *Trends in cell biology*, vol. 17, no. 3, pp. 127–134, 2007.

[23] D. Olins and A. Olins, "Chromatin history: Our view from the bridge," *Nat rev mol cell biol*, vol. 4, no. 10, pp. 809–814, 2003.

[24] K. Luger and T. Richmond, "Dna binding within the nucleosome core," *Current opinion in structural biology*, vol. 8, no. 1, pp. 33–40, 1998.

[25] K Ando and Y Kase, "Biological characteristics of carbon-ion therapy," *Int j radiat biol*, vol. 85, no. 9, pp. 715–728, 2009.

[26] G Barendsen, H Walter, J Fowler, and D Bewley, "Effects of different ionizing radiations on human cells in tissue culture: Iii. experiments with cyclotron-accelerated alpha-particles and deuterons," *Radiat res*, vol. 18, pp. 106–119, 1962.

[27] G Barendsen, "Dose fractionation, dose rate, and isoeffect relationships for normal tissue responses," *Int j radiat oncol biol phys*, vol. 8, pp. 1981–1997, 1982.

[28] J Woollard, T Blue, N Gupta, and R Gahbauer, "Development and calculation of an energy dependent normal brain tissue neutron rbe for evaluating neutron fields for bnct," *Health physics*, vol. 80, no. 6, pp. 583–589, 2001.

[29] N. Metropolis and S. Ulam, "The monte carlo method," *Journal of the american statistical association*, vol. 44, no. 247, pp. 335–341, 1949.

[30] S Agostinelli and et al., "Geant4- a simulation toolkit," *Nucl instr & methods in phys res a*, vol. A, no. 506, pp. 250–303, 2003.

[31] I El Naqa, P Pater, and J Seuntjens, "Monte carlo role in radiobiological modelling of radiotherapy outcomes," *Phys med biol*, vol. 57, R75–R97, 2012.

[32] M Bernal and J Liendo, "An investigation on the capabilities of the penelope mc code in nanodosimetry," *Med phys*, vol. 36, no. 2, pp. 620–625, 2009.

[33] C. Champion, S. Incerti, Y. Perrot, R. Delorme, M. Bordage, M. Bardiès, B. Mascialino, H. Tran, V. Ivanchenko, M. Bernal, Z. Francis, J.-E. Groetz, M. Fromm, and L. Campos, "Dose point kernels in liquid water: An intra-comparison between geant4-dna and a variety of monte carlo codes," *Applied radiation and isotopes*, vol. 83, Part B, pp. 137 –141, 2014, Quantum scattering codes and Monte Carlo simulations to model dynamical processes in biosystems.

[34] Z.Francis, S. Incerti, and et al., "Monte carlo simulation of energy-deposit clustering for ions of the same let in liquid water," *Phys med biol*, vol. 57, pp. 209–224, 2012.

[35] Z Francis, S Incerti, V Ivanchenko, C Champion, M Karamitros, M. A. Bernal, and Z. E. Bitar, "Monte carlo simulation of energy-deposit clustering for ions of the same let in liquid water," *Physics in medicine and biology*, vol. 57, pp. 209–224, 2012.

[36] T. André and et al., "Comparison of geant4-dna simulation of s-values with other monte carlo codes," *Nucl instr & methods in phys res b*, vol. 319, pp. 87–94, 2014.

[37] S Incerti, M Psaltaki, and et al., "Simulating radial dose of ion tracks in liquid water simulated with geant4-dna: A comparative study," *Nucl instr & methods in phys res b*, vol. 333, pp. 92–98, 2014.

[38] I Plante, A Ponomarev, and F Cucinotta, "3d visualisation of the stochastic patterns of the radial dose in nano-volumes by a monte carlo simulation of hze ion track structure," *Radiat protection dosimetry*, pp. 1–6, 2011.

[39] L Rapkin and et al., "A view of the chromatin landscape," *Micron*, vol. 43, pp. 150–158, 2012.

[40] E Lubeck and L Cai, "Single-cell systems biology by super-resolution imaging and combinatorial labeling," *Nature methods*, vol. 9, no. 7, pp. 743–750, 2012.

[41] K Hofer, X Lin, and M Schneiderman, "Paradoxical effects of iodine-125 decays in parent and daughter dna: A new target model for radiation damage," *Radiation research*, vol. 153, pp. 428–435, 2000.

[42] J Aten and et al., "Dynamics of dna double-strand breaks revealed by clustering of damaged chromosome domains," *Science*, vol. 303, pp. 92–95, 2004.

[43] M Falk, E Lukasova, and S Kozubek, "Chromatin structure influences the sensitivity of dna to gamma radiation," *Biochimica et biophysics acta*, vol. 1783, no. 12, pp. 2398–2414, 2008.

[44] C. K. Wang and X Zhang, "A nanodosimetry-based linear-quadratic model of cell survival for mixed-let radiations," *Physics in medicine and biology*, vol. 51, pp. 6087–6098, 2006.

[45] D Brenner, "The linear-quadratic model is an appropriate methodology for determining isoeffective doses at large doses per fraction," *Seminars in radiat oncol*, vol. 18, pp. 234–239, 2008.

[46] C Park, L Papiez, S Zhang, M Story, and R Timmerman, "Universal survival curve and single fraction equivalent dose: Useful tools in understanding potency of ablative radiotherapy," *Int j radiat oncol biol phys*, vol. 70, no. 3, 2008.

[47] M Guerrero and X Li, "Extending the linear-quadratic model for large fraction doses pertinent to stereotactic radiotherapy," *Phys med biol*, 2004.

[48] D Bettega, P Calzolari, and et al., "Neoplastic transformation induced by carbon ions," *Int j radiat oncol biol phys*, vol. 73, no. 3, pp. 861–868, 2009.

[49] "Relative biological effectiveness in ion beam therapy," IAEA and ICRU, Vienna, Technical Report 461, 2008.

[50] A Kellerer and H Rossi, "Rbe and the primary mechanism of radiation action," *Radiat res*, vol. 47, no. 1, pp. 15–34, 1971.

[51] E Hall, J Novak, and et al., "Rbe as a function of neutron energy: I. experimental observations," *Radiat res*, vol. 64, no. 2, pp. 245–255, 1975.

[52] G Barendsen, *Response of cultured cells, tumors, and normal tissues to radiations of different linear energy transfer*, M. Ebert and A. Howard, Eds., New York, 1968.

[53] M Belli, F Cera, and et al., "Rbe-let relationship for cell inactivation and mutation induced by low energy protons in v79 cells: Further results at the lnl facility," *Int j radiat biol*, vol. 74, no. 4, pp. 501–509, 1998.

[54] T Blue, N Gupta, and J Woollard, "A calculation of the energy dependence of the rbe of neutrons," *Phys med biol*, vol. 38, pp. 1693–1712, 1993.

[55] I Gibbs, C Patil, and et al., "Delayed radiation-induced myelopathy after spinal radiosurgery," *Neurosurgery*, vol. 64, no. 2, 2009.

[56] M. R. Raju, "Review: Proton radiobiology, radiosurgery, and radiotherapy," *Int j radiat biol*, vol. 67, no. 3, pp. 217–259, 1995.

[57] F. H. Attix, *Introduction to radiological physics and radiation dosimetry*. Weinheim: Wiley-VCH, 2004.

[58] M Bellamy, "A deterministic model of the relative biological effectiveness of photons and electrons," PhD thesis, Georgia Institute of Technology, Atlanta, GA, 2013.

[59] J Bedford and J Mitchell, "Dose-rate effects in synchronous mammalian cells in culture," *Radiat res*, vol. 54, no. 2, pp. 316–327, 1973.

[60] B Archer, J Gray, and et al., "Report no. 147 - structural shielding design for medical x-ray imaging facilities," National Council on Radiation Protection and Measurements, Tech. Rep., 2004.

[61] J Deye, J Rodgers, and et al., "Structural shielding design and evaluation for megavoltage x- and gamma-ray radiotherapy facilities," National Council on Radiation Protection and Measurements, Tech. Rep., 2005.

[62] C Karzmark, C Nunan, and E Tanabe, *Medical electron accelerators*. McGraw Hill, 1993.

[63] M Murphy, J Balter, and et al., "The management of imaging dose during image-guided radiotherapy: Report of the aapm task group 75," *Med phys*, vol. 34, no. 10, 2007.

[64] J. V. Dyk, J Galvin, G Glasgow, and E Podgorsak, "The physical aspects of total and half body photon irradiation," American Association of Physicists in Medicine, Tech. Rep., 1986.

[65] B Chang and R Timmerman, "Stereotactic body radiation therapy: A comprehensive review," *Amer j clin onc*, vol. 30, no. 6, 2007.

[66] M Schell, F Bova, and et al., "Stereotactic radiosurgery: Report of task group 42 radiation therapy committee," American Association of Physicists in Medicine, Tech. Rep., 1995.

[67] S Benedict, K Yenice, and et al., "Stereotactic body radiation therapy: The report of aapm task group 101," *Med phys*, vol. 37, no. 8, 2010.

[68] H Kubo, G Glasgow, and et al., "High dose-rate brachytherapy treatment delivery: Report of the aapm radiation therapy committee task group no. 59," *Med phys*, vol. 25, no. 4, 1998.

[69] S Ryu and et al., "Rtog 0631: Phase ii/iii study of image-guided radiosurgery/sbrt for localized spine metastasis," Radiation Therapy Oncology Group, Tech. Rep., 2011.

[70] R Timmerman and et al., "Rtog 618: A phase ii trial of stereotactic body radiation therapy (sbrt) in the treatment of patients with operable stage i/ii non-small cell lung cancer," Radiation Therapy Oncology Group, Tech. Rep., 2008.

[71] G Ezzell, J Burmeister, and et al., "Imrt commissioning: Multiple institution planning and dosimetry comparisons, a report from aapm task group 119," *Med phys*, vol. 36, no. 11, 2009.

[72] D Goodhead, "Inactivation and mutation of cultured mammalian cells by aluminum characteristic ultrasoft x-rays. iii. implication for theory of dual radiation action," *Int j radiat biol rel stud phys chem med.*, vol. 32, no. 1, pp. 43–70, 1977.

[73] M Hill and et al., "Biological effectiveness of isolated short electron tracks: V79-4 cell inactivation following low dose-rate irradiation with al-k ultrasoft x-rays," *Int j radiat biol*, vol. 78, no. 11, pp. 967–979, 2002.

[74] R Taleei and H Nikjoo, "The non-homologous end-joining (nhej) pathway for the repair of dna double-strand breaks: I. a mathematical model," *Radiation research*, vol. 179, no. 5, pp. 530–539, 2013.

[75] T Friedrich, M Durante, and M Scholz, "Modeling cell survival after irradiation with ultrasoft x rays using the giant loop binary lesion model," *Radiat res*, vol. 181, no. 5, pp. 485–494, 2014.

[76] C Griffin, D Stevens, and J. R. Savage, "Ultrasoft 1.5 kev aluminum k x rays are efficient producers of complex chromosome exchange aberrations as revealed by fluorescence in situe hybridization," *Radiat res*, 1996.

[77] R Nath, L Anderson, and et al., "Dosimetry of interstitial brachytherapy sources: Recommendations of the aapm radiation therapy committee task group no. 43," *Med phys*, vol. 22, no. 2, 1994.

[78] T Blue, J Woollard, N Gupta, and J Greskovich, "An expression for the rbe of neutrons as a function of neutron energy," *Phys med biol*, vol. 40, pp. 757–767, 1995.

[79] Y Maruyama, J Feola, C Magura, and J. L. Beach, "Relative biological effectiveness (r.b.e.) of cf-252 vs. acute co-60 and low dose rate cs-137 irradiation by spleen weight loss," *Radiat res*, vol. 27, pp. 112–120, 1986.

[80] C. K. Wang, X Zhang, and et al., "Experimental validation of the new nanodosimetry-based cell survival model for mixed neutron and gamma-ray radiation," *Phys med biol*, vol. 52, N367–N374, 2007.

[81] E Fortune, I Gauld, and C. K. Wang, "Gamma dose rate near a new cf-252 brachytherapy source," *Nucl tech*, vol. 175, pp. 73–76, 2011.

[82] J Watson and F Crick, "Molecular structure of nucleic acids: A structure for deoxyribose nucleic acid," *Nature*, vol. 171, pp. 737–738, 1953.

[83] J. R. Savage, "Insight into sites," *Mutat res*, 1996.

[84] Y Li and F Cucinotta, "Modeling non-homologous end joining," *Journal of theoretical biology*, vol. 283, pp. 122–135, 2011.

[85] W Purves and et al., *Life: The science of biology*, Seventh. Sinauer Associates and W.H. Freeman, 2004.

[86] E Lieberman-Aiden, N. L. van Berkum, and et al., "Comprehensive mapping of long range interactions reveals folding principles of the human genome," *Science*, vol. 326, no. 5950, pp. 289–293, 2009.

[87] A Miele and J Dekker, "Mapping cis- and trans- chromatin interaction networks using chromosome conformation capture (3c)," *Methods mol biol*, vol. 464, 2009.

[88] P Fraser and W Bickmore, "Nuclear organization of the genome and the potential for gene regulation," *Nature*, vol. 447, no. 24, pp. 413–417, 2007.

[89] T Cremer and C Cremer, "Chromosome territories, nuclear architecture and gene regulation in mammalian cells," *Nature*, vol. 2, pp. 292–301, 2001.

[90] R Taleei, P Girard, S Krishnaswami, and H Nikjoo, "The non-homologous end-joining (nhej) pathway for the repair of dna double-strand breaks: Ii. application to damage induced by ultrasoft x rays and low-energy electrons," *Radiation research*, vol. 179, no. 5, pp. 540–548, 2013.

[91] M Frankenberg-Schwager, A Gebauer, and et al., "Single-strand annealing, conservative homologous recombination, nonhomologous dna end joining, and the cell cycle-dependent repair of dna double-strand breaks induced by sparsely or densely ionizing radiation," *Radiat res*, vol. 171, pp. 265–273, 2009.

[92] C. K. Wang, "The progress of radiobiological models in modern radiotherapy with emphasis on the uncertainty issue," *Mutation research*, vol. 704, pp. 175–181, 2010.

[93] F Cucinotta, H Nikjoo, and D Goodhead, "Model for radial dependence of frequency distributions for energy imparted in nanometer volumes from hze particles," *Radiat res*, vol. 153, no. 4, pp. 459–468, 2000.

[94] A Ponomarev, S Costes, and F Cucinotta, "Stochastic properties of radiation-induced dsb: Dsb distributions in large scale chromatin loops, the hprt gene and within the visible volumes of dna repair foci," *Int j radiat biol*, vol. 84, no. 11, pp. 916–929, 2008.

[95] M. R. Raju and et al., "Radiobiology of ultrasoft x rays: I. cultured hamster cells (v79)," *Radiat res*, vol. 110, no. 3, pp. 396–412, 1987.

[96] D Alloni, A Campa, W Friedland, L Mariotti, and A Ottlenghi, "Integration of monte carlo simulations with pfge experimental data yields constant rbe of 2.3 for dna double-strand break induction by nitrogen ions between 125 and 225 kev/um let," *Radiat res*, vol. 179, pp. 690–697, 2013.

[97] J Ziegler and J Biersack, *Srim - the stopping and range of ions in solids*. Pergamon Press, 1985.

[98] C Lanctôt, T Cheutin, M Cremer, G Cavalli, and T Cremer, "Dynamic genome architecture in the nuclear space: Regulation of gene expression in three dimensions," *Nature*, vol. 8, pp. 104–115, 2007.

[99] D Peplow, "Direction cosines and polarization vectors for monte carlo photon scattering," *Nucl sci engr*, vol. 131, pp. 132–136, 1999.

[100] P Sundqvist, *Cylinder between 2 points*, 2004.

[101] S Holcombe, *Inpolyhedron - are points inside a triangulated volume - matlab central file exchange*, 2014.

[102] M Ester, H Kriegel, J Sander, and X Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," AAAI Press, 1996, pp. 226–231.

[103] E. W. Webster and et al., "A primer on low-level ionizing radiation and its biological effects," American Association of Physicists in Medicine, Tech. Rep., 1986.